

Department of Chemical and  
Environmental Engineering



**University of  
Nottingham**

UK | CHINA | MALAYSIA

# **Engineering of artificial metalloenzymes based on alcohol dehydrogenase scaffold**

Floriane Ludivine Martins

A thesis submitted to the University of Nottingham

For the degree of Doctor of Philosophy

March 2021

# Acknowledgements

Firstly, I am extremely grateful for the assistance and guidance provided by my supervisors Dr. Christof Jaeger and Dr. Anca Pordea. None of this thesis work would have been accomplished without their continuous advises, their patience throughout the thesis correcting phase, and their support in and outside work. These challenging PhD years were made achievable thanks to them. I would also like to extend my thanks to Prof. Rachel Gomes, for her advice during insightful discussions we had throughout the course of this project.

I will also give a special thanks to the Pordea group members: Azie, Camille, Candice, Harry, Mattias, Maria, Ricardo and Simone, for their help in the lab and also for all the fun we had outside.

Besides, I am grateful to Amy, Rachael and Steve for their technical support during my time in the lab and the knowledge sharing with the machines.

I would like to thank all the members of the SPT group, present and past: Andy, Camille, Charlotte, Clemency, Juliana, Lucas, Marko, Reza, Sophie, Victoria... It is their help, their sympathy, and the laughs we shared on many occasions that made my time in and outside the lab pleasant. They helped me to cope with hard aspects of the work.

To my friends I met in Nottingham I would like to say a big thanks for sharing those 4 fantastic years in the UK. Especially, thank you to Clemency and Aurélie who hosted me during the tough times of the pandemic, with plenty of fun, climbing and plants. I extend my gratitude to all my friends back in France, who came visiting me and contributed to clear my mind with the great time we spent discovering the UK.

Finally, I would like to express a massive thank you to my family. To my partner, Victor, for his support and confidence in me all along my PhD, especially in the tough times. I had a wonderful four years in the UK with you and counting. To my wonderful brother and parents, without their constant assistance during my entire academic career it would have been impossible to complete my studies.

# Abstract

An ongoing objective of the chemical industry is to reduce the use of eco-destructive chemicals in cost effective and sustainable synthesis routes for high value chemicals. Biocatalysis is an established strategy for green chemistry. In this approach, artificial metalloenzymes have emerged as advanced biocatalysts, able to support a wide range of non-naturally occurring reactions in a biological environment. They are built by combining the selectivity of enzymes with the appropriate non-natural reactivity of transition metal catalysts. Current examples of these hybrids often lack control over the structure and function. They are also based on proteins with no naturally evolved binding pockets, thus lacking the advantage of proximity between a wide range of substrates and the metal catalysts.

The research presented here investigates the design and development of artificial metalloenzymes for the transfer hydrogenation of imines, using the cofactor binding pocket of alcohol dehydrogenases to supramolecularly bind catalytic complexes.

A computational methodology was first created to conceptualise artificial metalloenzymes, starting with two NAD(P)H-dependent enzymes: the horse liver alcohol dehydrogenase (HLADH) and the *Thermoanaerobacter brockii* alcohol dehydrogenase (TbADH). The *in silico* study was used to understand the cofactor binding site and to identify the strongest non-covalent interactions. Computationally designed library of NADH analogues were then screened as anchors for the metal catalyst, resulting in the selection of a set of lead structures.

In the next step, a divergent total synthesis was proposed to create a small library of catalysts complexes *via* one key intermediate compound. Three iridium(III) catalysts were synthesised, based on the lead structures selected for their predicted high affinity in alcohol dehydrogenases.

The complexes were then tested for their capacity to form artificial metalloenzymes. The binding affinity was investigated *via* competition and ITC experiments, while the catalytic activity was assessed for the transfer hydrogenation of aromatic imines. This resulted in mutual inhibition of the metal complexes and the two wild type alcohol dehydrogenases, while the TbADH mutant without catalytic zinc ion formed a working artificial metalloenzyme.

# Table of contents

Acknowledgements.....	i
Abstract .....	ii
Table of contents .....	iii
List of Figures .....	vii
List of Tables.....	xviii
List of Schemes.....	xx
List of Abbreviations.....	xxiii
1 Literature review .....	1
1.1 Introduction.....	1
1.2 Artificial metalloenzymes: a strategy for biocatalysis.....	2
1.2.1 Background .....	2
1.2.2 Metal anchoring strategies .....	3
1.3 Alcohol dehydrogenases .....	21
1.3.1 Background .....	21
1.3.2 Horse liver alcohol dehydrogenase (HLADH) .....	22
1.3.3 Thermoanaerobacter brockii alcohol dehydrogenase (TbADH) .....	26
1.3.4 Cofactor analogues .....	29
1.3.5 Conclusion .....	31
1.4 Natural and synthetic hydrogen transfer reactions .....	33
1.4.1 Background .....	33



1.4.2	Chemical catalysis: Organocatalysts and metal catalysts .....	33
1.4.3	Biocatalysis: Transition metal catalysts within a protein scaffold host	36
1.5	Enzyme engineering .....	40
1.5.1	Background .....	40
1.5.2	Combination of rational design and directed evolution.....	41
1.5.3	Application to AMs engineering.....	42
1.5.4	Computational design of AMs.....	43
1.5.5	Conclusion .....	53
2	Aim and objectives.....	54
2.1	Research overall aim .....	54
2.2	Research objectives.....	55
2.2.1	In silico design of supramolecular anchors for transition metal catalysts	55
2.2.2	Chemical synthesis of the designed transition metal catalysts .....	55
2.2.3	Experimental creation and functionality testing of a designed AM ....	56
3	Docking guided design of artificial metal catalysts with affinity for alcohol dehydrogenases. ....	57
3.1	Introduction.....	57
3.2	Computational methods.....	57
3.2.1	Molecular Docking .....	57
3.2.2	Induced Fit Docking (IFD) calculations .....	62
3.3	Results and discussion.....	62

3.3.1	Systematic docking evaluations to study the cofactor binding pocket	63
3.3.2	Docking with HLADH co-crystallised with substrate .....	76
3.3.3	Libraries docking .....	81
3.3.4	Conclusion of the first docking study in HLADH .....	87
3.3.5	Improved docking strategy and design of small anchor library .....	88
3.3.6	Comparative study with TbADH .....	107
3.4	Conclusion and selection of hit structures .....	111
4	Synthesis of metal-based catalysts for the supramolecular anchoring of iridium transition metal catalysts inside ADHs .....	113
4.1	Background .....	113
4.2	Results and discussion .....	117
4.2.1	Synthesis of the ketone-substituted intermediate .....	117
4.2.2	Synthesis of the key intermediate $\beta$ -diketoester .....	118
4.2.3	Transition metal complexation of all deprotected compounds .....	139
4.3	Conclusion .....	141
4.4	Experimental methods .....	142
4.4.1	Materials .....	142
4.4.2	Methods .....	142
5	Enzyme engineering: non-covalent conjugation and catalysis of imine reduction	159
5.1	Introduction .....	159
5.2	Materials and Methods .....	160

5.2.1	Reagents and materials.....	160
5.2.2	Methods .....	161
5.3	Results and discussion.....	168
5.3.1	Production of functional enzymes: His-HLADH, TbADH-WT and TbADH-5M	168
5.3.2	Investigation of the binding affinity between the transition metal catalysts and the enzymes.....	170
5.3.3	Assessing the catalytic reactivity of the three metal catalysts and the AMs	180
5.4	Conclusion .....	188
6	Discussion, conclusions and perspectives .....	190
7	References .....	197
8	Appendices.....	222

## List of Figures

<b>Figure 1:</b> Representation of metal anchoring methods, to build AMs. The following codes were applied: the amino acid (Aa), the covalent or supramolecular ligands and spacers (L), the ligand coordinating the metal ion (red crescent), the metal ion catalyst (red ball). The dashed lines represent non-covalent interactions, and the plain lines represent covalent interactions.....	3
<b>Figure 2:</b> Representation of the LmrR engineered <i>in vivo</i> with incorporation of BpyAla residues to create AM. After coordination of the Cu(II) ion, an example of its new to nature catalysis of Friedel-Crafts alkylation reaction is presented. Adapted from Roefle article. <sup>23</sup> .....	6
<b>Figure 3:</b> Reproduction of Baker and co-workers designed protein. <sup>25</sup> <b>A)</b> Solved crystal structure of their designed protein with Co <sup>2+</sup> <b>B)</b> and Ni <sup>2+</sup> . ....	6
<b>Figure 4:</b> Representation of ligands structures used in literature to make covalent linkage in order to anchor metal ions inside a host to form AMs. ....	8
<b>Figure 5:</b> Reaction scheme of the bioconjugation of lipase cutinase hybrid, adapted from Gebbink and co-worker. <sup>37</sup> .....	9
<b>Figure 6:</b> MnSalen complex inside apo-Mb protein scaffold, reproduced from Lu and co-workers. <sup>40</sup> <b>A)</b> MnSalen complex <b>B)</b> Computer model of MnSalen dual anchoring in Mb, overlayed with the protein natural heme cofactor. ....	9
<b>Figure 7:</b> Example of Cr(III)-salophen catalyst complex, reported from Watanabe and co-workers. <sup>45</sup> .....	11
<b>Figure 8:</b> C-H amination reaction scheme catalysed by the best Ir(Me)-PIX CYP119 enzyme variant. Figure adapted from Hartwig and co-workers. <sup>47</sup> .....	12
<b>Figure 9:</b> Representation of supramolecular assembly of Cu(II)-phenantroline complex catalyst inside LmrR, with an example of its new to nature Friedel-Crafts alkylation of indole. Adapted from Roelfes and co-workers. <sup>23</sup> .....	12
<b>Figure 10:</b> Presentation of the first AM based on Av/Sav technology, catalysing the hydrogenation of $\alpha$ -acetamidoacrylic acid. This AM was composed of a biotinylated Rh catalyst non-covalently bound inside avidin. Figure adapted from Wilson and Whitesides. <sup>50</sup> .....	13

<b>Figure 11:</b> Examples of biotinylated metal complexes, designed to be non-covalently embedded inside Sav to build artificial hydrogenases. [Rh(COD)Biot-1] is constructed with the first generation of diphosphine ligand. The other complexes were then optimised with amino acid spacers. <sup>49,55</sup> .....	14
<b>Figure 12:</b> AM model for the enantioselective transfer hydrogenation of carbonyl, adapted from Ward and co-workers. <sup>58</sup> The ligand biotin (blue) is separated from the d <sup>6</sup> piano-stool metal complex (red) by different spacers (green). The prochiral substrate receives two hydrogens without coordinating with the metal. ....	15
<b>Figure 13:</b> An example the asymmetrical reduction from cyclic imine to amines, catalyse by an AM based on the biotin-Sav technology. ....	16
<b>Figure 14:</b> Enantioselective benzannulation catalysed by a new Sav-AM, adapted from Ward and co-workers. <sup>69</sup> <b>A)</b> Synergistic action of a basic residue introduced by site-directed mutagenesis with a biotinylated RhCp*Cl catalyst non-covalently bound to Sav. <b>B)</b> Reaction scheme of the synthesis of dihydroisoquinolones by benzannulation reaction catalysed by the new Sav AM. ....	17
<b>Figure 15:</b> Representation of the double deracemization with MAO/artificial imine reductase (ATHase)/catalase, reproduced from Ward and co-workers. <sup>70</sup> .....	18
<b>Figure 16:</b> Two AMs adapted from Ward and co-workers. <b>A)</b> AM for the transfer hydrogenation of salsolidine. The bidentate ligand (blue) anchor the IrCp* metal complex inside hCAII. <sup>83</sup> <b>B)</b> AM for ring-closing metathesis. The arylsulfonamide ligand (blue) is anchoring the Grubbs-Hoveyda type catalyst inside hCAII. <sup>85</sup> .....	20
<b>Figure 17:</b> Horse liver ADH (HLADH) structure (PDB 4xd2) with NAD(H) cofactor (green) at the domain's interfaces. Zn <sup>2+</sup> are displayed in red spheres. <sup>92</sup> .....	22
<b>Figure 18:</b> Catalytic mechanism of HLADH based on proton-relay, reproduced from Lee and co-workers. <sup>93</sup> .....	23
<b>Figure 19:</b> Structure of the ADHs natural cofactors <b>A)</b> NADH and <b>B)</b> NADPH. The structure is divided in two parts: the adenine ring and the nicotinamide ring, used for the catalytic activity. The two parts are linked by a diphosphate group. Their oxidized forms are respectively NAD <sup>+</sup> and NADP <sup>+</sup> . ....	24
<b>Figure 20:</b> Schematic representation of binding interactions between HLADH crystal structure (PDB 4xd2) and NAD <sup>+</sup> : amide binding (blue), nicotinamide ribose binding (green), pyrophosphate binding (red) and adenine ribose binding (orange). ....	25

<b>Figure 21:</b> Holo-TbADH crystal structure at 2.5 Å (PDB 1YKF) with its natural cofactor (green) and the Zn <sup>2+</sup> ions in red spheres. <sup>97</sup> .....	26
<b>Figure 22:</b> Reproduction of the catalytic mechanism of TbADH from Lee and co-workers. <b>Int.1</b> and <b>Int.2</b> are the two intermediates penta-coordinated. <sup>93</sup> .....	27
<b>Figure 23:</b> Interactions of TbADH (PDB 1ykf) with NADPH: amide binding (blue), nicotinamide ribose binding (green), pyrophosphate binding (red) and adenine phosphate ribose binding (orange). <sup>97</sup> .....	29
<b>Figure 24:</b> CL4 mimic is an analogue of NADH cofactor, designed by Ansell and co-workers. Only the nicotinamide ring was preserved and substituted by triazines. <sup>103</sup> .....	30
<b>Figure 25:</b> 1-Benzylnicotinamide, used by Fish and co-workers to mimic NADH cofactor. ....	30
<b>Figure 26:</b> NADH mimics reduced form, tested inside ER enzymes by Hollman and co-workers for oxidoreduction reactions. <sup>107</sup> .....	31
<b>Figure 27:</b> Asymmetric transfer hydrogenation of carbonyl, catalysed by Noyori's chiral Ru complexes. The metal complex favoured transition state is depicted underneath the reaction scheme. Adapted from Noyori's work. <sup>118</sup> .....	34
<b>Figure 28:</b> Hypothetic catalytic mechanism for the enantioselective transfer hydrogenation of imine, using a biotinylated Ru bidentate catalyst in aqueous solution with sodium formate as hydride source. Adapted from Suss-Fink and co-workers postulated mechanism. <sup>124</sup> .....	35
<b>Figure 29:</b> Transfer hydrogenation of double bonded carbon, catalysed by iminium combined with Hantzsch ester, separately reported by List and MacMillan groups. <sup>128</sup> .....	36
<b>Figure 30:</b> Artificial transfer hydrogenase created by Duhme-Klair and co-workers, combining CeuE protein and a sulfonamide-Ir(II) complex. <b>A)</b> Catalyst engineered for the synthesis of salsolidine. In blue is highlighted the iron-siderophore ligand, and in red the Ir-complex catalyst. <b>B)</b> Design of the siderophore (blue) anchoring the iridium catalyst (red) inside CeuE. Adapted from Duhme-Klair and co-workers <sup>145</sup> .....	39
<b>Figure 31:</b> Scheme of two approaches of protein engineering: rational design and directed evolution. Adapted from Gruber-Khadjawi and co-workers representation. <sup>147</sup> .....	40

<b>Figure 32:</b> Combination of Rational Design and Directed: studied enzyme structures and properties are used to select hot spot amino acids (rational design). These are replaced by all possible amino acids to create a smaller and efficient library (directed evolution). The library is screened, and the best variant is selected. Adapted from Gruber-Khadjawi and co-workers representation. <sup>147</sup> .....	42
<b>Figure 33:</b> DeGrado <i>de novo</i> proto-enzymes. <b>A)</b> Top: Proposed mechanism of the catechol oxidation reaction catalysed by DF3. Bottom: Representation of a four chain DF structure (similar to DF3). <b>B)</b> X-ray structure of PS1 (PDB 5WLM) with a close view of the zinc cluster. Figures adapted from DeGrado and co-workers. <sup>189</sup> .....	53
<b>Figure 34:</b> Iridium coordinated with chloride, bidentate sulphonamide ligand and Cp*. This transition metal complex was built using the Maestro building panel.....	61
<b>Figure 35:</b> Established docking protocol used to filter various ligands in order to select lead structures. The two Glide dockings give a rough ranking of structures, followed by a refinement using MM-GBSA method.....	63
<b>Figure 36:</b> NADH (blue) re-docked inside HLADH and superposed to the co-crystallised cofactor (green). Hydrogen bonding contacts are represented in dashed lines and the amino acids in contact with the ligands are labelled and shown in pink. ....	65
<b>Figure 37:</b> <b>A)</b> Superposition of the crystal nicotinamide ring of NADH (yellow) and the structure from QM calculations (cyan) from Meijers and co-workers <b>B)</b> Superposition of NADPH cofactor (blue) with the NADH (green) from docking experiment. The highlighted D223 repulsed the NADPH ribose phosphate from the adenine group. By contrast, the ribose of NADH made contacts with D223.....	67
<b>Figure 38:</b> Fragments of NAD(P)H structure based on the nicotinamide and adenosine parts, designed for the docking inside HLADH.....	68
<b>Figure 39:</b> Nicotinamide mononucleotide fragment (NMN, fragment 6, blue) superposed to NADH (green). The H-bonds are in dashed yellow lines, while the amino acids (pink) in contact with the ligands are labelled. ....	70
<b>Figure 40:</b> The nicotinamide fragment 7 (blue) superposed with NADH (green), slide inside the hydrophobic substrate binding pocket. ....	70
<b>Figure 41:</b> Example of adenosine monophosphate fragment 9 (blue) superposed to NADH (green), and its interaction residues (pink). ....	71

**Figure 42:** Designed ligands for further docking screening. Structures **1**, **15** and **16** are based on previous work from literature.<sup>104,109</sup> Structures **17** to **22** are based on NMN fragments, supplemented with the sulphonamide bidentate ligand for structures **19** to **22**. ..... 72

**Figure 43:** Best scored docking poses of ligand **15** and **16** (blue) superposed to NADH (green). The benzyl was inserted in the substrate hydrophobic pocket. The ligands bound to valines (pink) with their amide and sulfonamide group. .... 74

**Figure 44:** 3D structure and corresponding interaction diagram of ligand **20** (blue). Phosphate and ribose made interactions that reversed the structure's orientation, inserting the sulphonamide group inside the substrate pocket, ahead of the NADH binding pocket. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, Ile: isoleucine, Leu: leucine, Lys: lysine, Met: methionine, Phe: phenylalanine, Pro: proline, Ser: serine, Thr: threonine, Val: valine. .... 75

**Figure 45: A)** Docking poses of ligands **18** and **20** (blue), compared to crystal structure **NAD<sup>+</sup> 2** (green), in presence of the substrate (red). **B)** Pose of ligand **20** (blue) superposed with **NAD<sup>+</sup> 2** (green). **C)** Interaction diagram represents ligand **20** contacts with the amino acids inside NADH binding pocket. The red box represents interaction similar to NADH pyrophosphate and the orange box the NADH adenine ribose interactions. Arg: arginine, Asp: aspartic acid, Asn: asparagine, Gly: glycine, Ile: isoleucine, Leu: leucine, Lys: lysine, Phe: phenylalanine, Pro: proline, Thr: threonine, Val: valine. .... 80

**Figure 46:** Interaction diagram of best MM-GBSA scoring ligand in HLADH cofactor binding pocket. Similar interactions as phosphate (red box), ribose (green) and nicotinamide (blue) groups of NADH were spotted. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Met: methionine, Phe: phenylalanine, Ser: serine, Thr: threonine, Val: valine. .... 81

**Figure 47:** Selection of the ligands poses from libraries with the best XP and MM-GBSA scores, and their corresponding interaction diagrams inside the cofactor binding pocket of HLADH. Ligands (blue) are superposed to **NAD<sup>+</sup>2** (green) with substrate (red). In the diagrams, the similar interactions as the NADH pyrophosphate group are represented in red boxes and the nicotinamide in blue. Arg: arginine, Asp:



aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Phe: phenylalanine, Ser: serine, Thr: threonine, Val: valine. ....	85
<b>Figure 48:</b> Pose and interaction diagram of the best scoring fragment from the online libraries (blue). The interactions mimicked NAD <sup>+</sup> ribose (green box) and phosphate interactions (red box). Arg: arginine, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Ser: serine, Thr: threonine, Val: valine. ....	86
<b>Figure 49:</b> An example of the lowest ranked structure from Schrodinger library, with few interactions. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Phe: phenylalanine, Pro: proline, Ser: serine, Thr: threonine, Val: valine.....	87
<b>Figure 50:</b> The protocol designed for the second set of docking experiments, with a first rough screening using Glide score and a refinement using IFD calculation with some enzyme flexibility in the final step. ....	89
<b>Figure 51:</b> The prototype of the new high affinity ligand structure was based on the nicotinamide part of NAD(P) <sup>+</sup> cofactor. The phosphate and ribose represent the anchoring part and can be replaced by shorter and simpler hydrophilic groups (blue). A sulphonamide bidentate metal ligand with and without its typical aromatic ring (red), is inserted at various positions on the nicotinamide ring. ....	91
<b>Figure 52:</b> Selection of the functional groups ( <b>R</b> ) composing the new library of ligands. They were created using variation of structure length and hydrophilic groups combinations.....	92
<b>Figure 53:</b> Model of ligand structures for the second docking screening, where <b>R</b> represents the various substituents (see <b>Figure 52</b> ) and both phenyl and pyridine rings were used. The sulphonamide served as support for the metal complex and was inserted in ortho, meta or para position in the aromatic ring.....	93
<b>Figure 54:</b> General process of the Glide SP / Glide XP / IFD created protocol and the selection of ligands structures at each step. ....	94
<b>Figure 55:</b> Superposition of <b>A</b> ) Anchor <b>R10</b> (blue) with XP score of -10.3 kcal mol <sup>-1</sup> in an efficient pose, with NADH (green) and <b>B</b> ) Anchor <b>R1</b> (blue) with XP score of -10.7 kcal mol <sup>-1</sup> in a reversed pose, with NADH (green). The sulphonamide would locate the metal toward the solvent.....	95

<b>Figure 56:</b> Model of ligands from the library previously created ( <b>R</b> in <b>Figure 52</b> ), in complex with the iridium transition metal. ....	95
<b>Figure 57:</b> Graph representing the final score (y axis) obtained from the affinity score and the structures poses score (orientation/location) after HLADH docking. All the designed ligands (x axis) are represented with various metal complex insertions (ortho, meta, para) and the C-ring mode (orange) or the N-ring (blue).....	98
<b>Figure 58:</b> <b>A)</b> Ligands <b>R5<sub>ortho</sub></b> (XP= -4.0 kcal mol <sup>-1</sup> , blue) with a wrong orientation for catalysis (superposed to NADH green). <b>B)</b> Ligand <b>R2<sub>ortho</sub></b> (XP= -6.1 kcal mol <sup>-1</sup> , blue) with a suitable pose toward the substrate (red). ....	99
<b>Figure 59:</b> Set of lead catalyst structures, selected based on the IFD results. The anchoring part of these products should be suitable for the anchoring of the iridium transition metal inside HLADH.....	100
<b>Figure 60:</b> Ligand <b>R25</b> , with a C-ring and a meta insertion of the metal complex. <b>A)</b> Superposition of <b>R25</b> (blue) with NAD <sup>+</sup> (green). <b>B)</b> Interaction diagram of <b>R25</b> inside HLADH cofactor pocket. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Ser: serine, Thr: threonine, Val: valine. ....	101
<b>Figure 61:</b> Ligand <b>R2</b> , with a C-ring and a meta insertion of the metal complex. <b>A)</b> Superposition of <b>R2</b> (blue) with NAD <sup>+</sup> (green). <b>B)</b> Interaction diagram of <b>R2</b> inside HLADH cofactor pocket. Arg: arginine, Asn: Asparagine, Asp: aspartic acid, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Ser: serine, Thr: threonine, Val: valine. ....	102
<b>Figure 62:</b> Ligand <b>R9</b> , with a C-ring and an ortho insertion of the metal complex. <b>A)</b> superposition of <b>R9</b> (blue) with NAD <sup>+</sup> (green). <b>B)</b> Interaction diagram of <b>R9</b> inside HLADH cofactor pocket. Arg: arginine, Ala: alanine, Glu: glutamic acid, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Phe: phenylalanine, Ser: serine, Thr: threonine, Val: valine. ....	103
<b>Figure 63:</b> Comparison of IFD scores between the same catalyst structures with either a N or C-ring. ....	104
<b>Figure 64:</b> The number of contacts compared between <b>A)</b> an anchors <b>R10</b> with only sulphonamide support and <b>B)</b> metal complex catalyst. ....	106

<b>Figure 65:</b> Truncated hit ligands used for enhanced affinity tests. They were composed only of an aromatic ring and the <b>R</b> anchoring part.....	106
<b>Figure 66:</b> Graph representing the final score (y axis) obtained from the affinity score and the structures poses score (orientation/location) by IFD calculations in TbADH. All the designed ligands (x axis) are represented with various metal complex insertions (ortho, meta, para) and the C-ring mode (orange) or the N-ring (blue). .....	110
<b>Figure 67:</b> <sup>1</sup> H-NMR spectrum comparison showing the proton shift differences between the two tautomers of the β-diketoester <b>31</b> : blue <b>31a</b> and red <b>31b</b> (full spectrum in Appendix 1). .....	123
<b>Figure 68: A)</b> H-H coupling in product <b>43</b> . The arrows represent the couplings made by the H11. <b>B)</b> The COSY-NMR spectrum of product <b>43</b> showed the couplings between H11 and H10 and between H11 and H13, highlighted by circles. ....	133
<b>Figure 69:</b> <sup>1</sup> H-NMR spectrum of product <b>44</b> . The three new observed aromatic signals correspond to furan hydrogens. ....	135
<b>Figure 70: A)</b> H-H coupling in product <b>46</b> . The blue arrows represent the couplings made by the H7. The red arrows represent the couplings made by the H10. <b>B)</b> The COSY-NMR spectrum of product <b>46</b> showed the coupling of the new H7 with the two neighbours H9, highlighted by a circle.....	139
<b>Figure 71:</b> By-product of the total reduction of <b>45</b> aromatic ketone, under Cbz deprotection conditions using Pd/C - H <sub>2</sub> .....	139
<b>Figure 72:</b> Schematic representation of AMs with TbADH-WT scaffold (left) and after site-directed mutagenesis with TbADH-5M lacking the catalytic Zn <sup>2+</sup> (right). The loss of the metal ion freed space for the transition metal catalyst.....	160
<b>Figure 73:</b> Example of initial rate of reaction for an enzymatic assay in the oxidative direction. NAD(P)H absorbance was monitored at 340 nm, with butan-2-ol and ethanol as the substrates: -●- TbADH-WT 150 mM butan-2-ol -●- His-HLADH 200 mM ethanol -●- Control. ....	164
<b>Figure 74:</b> Protein SDS-PAGE of three ADHs variants after purification of <b>A)</b> TbADH-WT, <b>B)</b> TbADH-5M and <b>C)</b> His-HLADH. F= flowthrough and L= Ladder.....	169
<b>Figure 75:</b> Representation of the c-value influence on the sigmoidal curve. When the slope is too shallow (c = 0.5) or too stipe (c = 5000), the K <sub>d</sub> can't be defined with accuracy. <sup>296</sup> .....	172

<b>Figure 76:</b> ITC experiment with the enzyme natural cofactor at [E]=100 $\mu$ M and [NAD(P)H] = 2100 $\mu$ M. <b>A)</b> In TbADH-5M, the left panel displays an example of raw data of the heat pulses resulting from each titration. On the right, the curve represents the integration of the heat, normalised per mole of NADPH; <b>B)</b> TbADH-WT and <b>C)</b> His-HLADH. The molar ratio refers to titrant (ligands) / titrand (enzymes).	173
<b>Figure 77:</b> ITC experiment using 2100 $\mu$ M NADP <sup>+</sup> and 100 $\mu$ M enzyme concentration. <b>A)</b> TbADH-5M and <b>B)</b> TbADH-WT. The molar ratio refers to titrant (ligands) / titrand (enzymes).	174
<b>Figure 78:</b> Example of ITC experiments at [E]= 100 $\mu$ M and [L]=2100 $\mu$ M with <b>A)</b> E= His-HLADH, L= acid <b>36</b> and <b>B)</b> E= TbADH-WT, L= $\beta$ -diketoacid <b>38</b> . The molar ratio refers to titrant (ligands) / titrand (enzymes).	175
<b>Figure 79:</b> ITC experiment with [E]=130 $\mu$ M and [L]=2400 $\mu$ M with <b>A)</b> E= TbADH-WT, L= $\beta$ -diketoacid <b>38</b> <b>B)</b> E= His-HLADH, L= $\beta$ -diketoacid <b>38</b> . The molar ratio refers to titrant (ligands) / titrand (enzymes).	175
<b>Figure 80:</b> Synthesised anchors and their corresponding iridium complexes used in this affinity study. Both types of compounds will be tested for their supramolecular binding strength and specificity within the wild type TbADH-WT and His-HLADH.	177
<b>Figure 81:</b> IC <sub>50</sub> values of synthesised anchors and iridium complexes with His-HLADH (blue) and TbADH-WT (orange) in the oxidative reaction. NAD(P) <sup>+</sup> (0.5 mM), ethanol (4 mM) and butan-2-ol (150 mM) were used with His-HLADH (15 $\mu$ g) and TbADH-WT (1 $\mu$ g) in 100 mM TrisHCl pH 8 buffer. The molecules ranged from 10 <sup>-4</sup> to 10 <sup>2</sup> mM (> 10mM for compound <b>46</b> : not enough stocks to test further).	178
<b>Figure 82:</b> The product and substrate concentrations obtained from the reduction of 4 mM cyclic imine at 38 °C in PBS buffer, catalysed by 50 $\mu$ M iridium complex <b>50</b> and 50 $\mu$ M enzymes, for 24h (25°C for His-HLADH reaction). The graph presents the iridium complex alone and in combination with the three selected enzymes for AMs engineering: His-HLADH, TbADH-WT and TbADH-5M.	182
<b>Figure 83:</b> The product and substrate concentrations obtained from the reduction of 4 mM cyclic imine at 38 °C in PBS buffer, for 24h. The graph presents the iridium complex <b>50</b> with TbADH-5M and TbADH-WT, at various concentrations combinations.	183

<b>Figure 84:</b> Proposed mechanism of the inhibition of $d^6$ piano-stool iridium complexes by binding to amino acids, adapted from Hollmann and co-workers theory. <sup>300</sup> .....	184
<b>Figure 85:</b> The product and substrate concentrations obtained from the reduction of 4 mM cyclic imine at 38 °C in PBS buffer, catalysed by iridium complex <b>49</b> alone and in combination with the three enzymes, for 24h (25°C for His-HLADH reaction)...	186
<b>Figure 86:</b> The product and substrate concentrations obtained from the reduction of 4 mM imine at 38 °C in PBS buffer, catalysed by 50 $\mu$ M iridium complex <b>47</b> alone and with 50 $\mu$ M of the three enzymes His-HLADH, TbADH-WT and TbADH-5M, for 24h (25°C for His-HLADH reaction). ....	187
<b>Figure 87:</b> Synthesised metal catalyst complexes: Ir-triol <b>50</b> , Ir-acid <b>47</b> and Ir-furan <b>49</b> . ....	192
<b>Figure 88:</b> Schematic representation of the rational design of an AM and the possible optimisations cycle.....	196
<b>Figure 89:</b> Michaelis-Menten curve of HLADH in the oxidation of ethanol. ....	256
<b>Figure 90:</b> Inhibition test on His-HLADH with anchor benzoic acid <b>36</b> .....	256
<b>Figure 91:</b> Inhibition test on TbADH-WT with anchor benzoic acid <b>36</b> . ....	257
<b>Figure 92:</b> Inhibition test on His-HLADH with anchor furan <b>44</b> . ....	257
<b>Figure 93:</b> Inhibition test on TbADH-WT with anchor furan <b>44</b> .....	257
<b>Figure 94:</b> Inhibition test on His-HLADH with anchor Triol <b>46</b> .....	258
<b>Figure 95:</b> Inhibition test on TbADH-WT with anchor Triol <b>46</b> . ....	258
<b>Figure 96:</b> Inhibition test on His-HLADH with anchor [IrClCp*]-benzoic acid <b>47</b> . ...	258
<b>Figure 97:</b> Inhibition test on TbADH-WT with anchor [IrClCp*]-benzoic acid <b>47</b> . ...	259
<b>Figure 98:</b> His-HLADH inhibition test with anchor [IrClCp*]-furan <b>49</b> .....	259
<b>Figure 99:</b> Inhibition test on TbADH-WT with anchor [IrClCp*]-furan <b>49</b> . ....	259
<b>Figure 100:</b> Inhibition test on His-HLADH with anchor [IrClCp*]-triol <b>50</b> .....	260
<b>Figure 101:</b> Inhibition test on TbADH-WT with anchor [IrClCp*]-triol <b>50</b> . ....	260
<b>Figure 102:</b> Calibration curve of the imine 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline substrate and the two chiral amine 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline products (Amine 1 and 2). ....	260
<b>Figure 103:</b> The kinetic profile for the reduction of 4mM 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline at 38 °C catalysed by 50 $\mu$ M iridium complex [IrClCp*]-benzoic acid <b>47</b> . ....	261

<b>Figure 104:</b> The kinetic profile for the reduction of 4mM 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline at 38 °C catalysed by 50 µM iridium complex [IrClCp*]-furan <b>49</b> . .....	261
<b>Figure 105:</b> The kinetic profile for the reduction of 4mM 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline at 38 °C catalysed by 50 µM iridium complex [IrClCp*]-triol <b>50</b> . .....	262

## List of Tables

<b>Table 1:</b> Cofactors docking scores represented by Glide XP (kcal mol <sup>-1</sup> ) and MM-GBSA calculations ( $\Delta G$ in kcal mol <sup>-1</sup> ). The RMSD values compare the dock poses to the co-crystallised cofactors in the X-ray structures.....	64
<b>Table 2:</b> Cofactor docking inside HLADH (PDB 4XD2). Glide XP score (kcal mol <sup>-1</sup> ) and the free binding energy $\Delta G$ (kcal mol <sup>-1</sup> ) from MM-GBSA calculations are presented for all cofactors. The lower the score, the better the binding affinity.....	66
<b>Table 3:</b> Glide XP (kcal mol <sup>-1</sup> ) and MM-GBSA ( $\Delta G$ in kcal mol <sup>-1</sup> ) results ranked by HLADH XP score. The highlighted numbers represent the best scores. The location of fragments compared to NADH is presented in the last column.....	69
<b>Table 4:</b> Presentation of Glide XP (kcal mol <sup>-1</sup> ) and MM-GBSA ( $\Delta G$ kcal mol <sup>-1</sup> ) docking scores for the designed compounds and NAD(P)H inside HLADH with and without Zn <sup>2+</sup> . The ligands are ranked by XP results in HLADH. The highlights represent the best scores. The location of ligand nicotinamide ring is compared to NADH for both docking inside HLADH and HLADH without Zn <sup>2+</sup> .....	73
<b>Table 5:</b> Glide XP (kcal mol <sup>-1</sup> ) and MM-GBSA ( $\Delta G$ in kcal mol <sup>-1</sup> ) results inside HLADH with substrate, ranked by XP score. The highlighted numbers represent the best scores. The location of fragments is compared to the co-crystallised NAD <sup>+</sup> in the last column. HLADHsubst refers to the crystal structure containing a substrate (PDB 4DWV). .....	77
<b>Table 6:</b> XP (kcal mol <sup>-1</sup> ) and MM-GBSA ( $\Delta G$ in kcal mol <sup>-1</sup> ) docking results, for designed ligand inside HLADH including co-crystallised substrate with and without Zn <sup>2+</sup> . The ligands are ranked by XP results and the highlights represent the best scores. The location of ligand nicotinamide ring is compared to the NAD <sup>+</sup> co-crystal. HLADHsubst refers to the crystal structure containing a substrate (PDB 4DWV). .....	78
<b>Table 7:</b> Glide XP results (kcal mol <sup>-1</sup> ) of fragment into HLADH, and their position compared to the natural cofactor.....	107
<b>Table 8:</b> Docking of NAD(P) cofactors in TbADH, following a docking protocol set in the previous section (3.3.1.1, page 62). The Glide XP scores are in kcal mol <sup>-1</sup> . .....	108

<b>Table 9:</b> Comparison of hit anchors in TbADH and both HLADH with and without substrate. The ranking of the structures is made by their filters score. XP score (kcal mol <sup>-1</sup> ) obtained from IFD calculations is also shown.....	110
<b>Table 10:</b> Claisen condensation reaction between ketone <b>30</b> and diethyl oxalate under various conditions. ....	119
<b>Table 11:</b> Different conditions applied for the Boc deprotection reaction of <b>31</b> ....	125
<b>Table 12:</b> Metalation of the deprotected compounds <b>36</b> , <b>37</b> , <b>44</b> and <b>46</b> using [Ir <sub>2</sub> Cl <sub>2</sub> Cp*] <sub>2</sub> and following the modified protocol from Letondor and co-workers. <sup>58</sup> .....	140
<b>Table 13:</b> Buffers used for His-HLADH purification. Buffer A corresponded to the binding buffer, buffer B was the elution buffer and buffer C the stock buffer. ....	162
<b>Table 14:</b> Buffers used for TbADH-WT and TbADH-5M purification. Buffer A corresponded to the binding buffer, buffer B was the elution buffer and buffer C the stock buffer. ....	163
<b>Table 15:</b> Final concentrations for the oxidative and reductive reactions of His-HLADH and TbADH-WT.....	165
<b>Table 16:</b> Final concentrations for oxidative Michaelis-Menten assay on His-HLADH. ....	165
<b>Table 17:</b> Final concentrations for oxidative inhibition assays.....	166
<b>Table 18:</b> Specific activity of His-HLADH and TbADH-WT for the oxidation and reduction of their natural substrates, compared to the literature reports. <sup>287,294</sup> One unit of activity is defined by the amount of enzyme that catalyses the formation of 1 μmol of NAD(P)H per minute (U / mg). The protocol is described in section 5.2.2.2, page 166.....	170
<b>Table 19:</b> Turnover frequency (TOF) of the three iridium catalysts is presented for the reduction of 4mM cyclic imine substrate with 50 μM of catalysts at 38°C after 1h of reaction. The experiments were run in triplicates for 24h and the results presented are the mean ± standard error.....	181



## List of Schemes

<b>Scheme 1:</b> Retrosynthesis designed to produce the previously selected hit metal catalysts structures. ....	114
<b>Scheme 2:</b> General synthetic pathway for the production of metal catalysts. Reagents and conditions: a) Boc <sub>2</sub> O, DCM; b) triethylamine, Boc-ethylenediamine, DCM; c) (i) SOCl <sub>2</sub> , CH <sub>3</sub> NHOCH <sub>3</sub> , Et(NPr- <i>i</i> ) <sub>2</sub> , dry DCM; (ii) CH <sub>3</sub> MgBr, dry THF; d) CH <sub>3</sub> Li, DEE; e) <i>t</i> -BuOLi, dry THF; f) (i) TFA, DCM, (ii) NaOH, MeOH/H <sub>2</sub> O, (iii) [IrCl <sub>2</sub> Cp*] <sub>2</sub> , Et <sub>3</sub> N, MeOH; g) (i) NaOH, MeOH/H <sub>2</sub> O, (ii) TFA, DCM h) (i) NaBH <sub>4</sub> , MeOH, (ii) [IrCl <sub>2</sub> Cp*] <sub>2</sub> , Et <sub>3</sub> N, MeOH; i) <i>t</i> -BuOLi, ethyl propionate, dry THF; j) (i) NaBH <sub>4</sub> , MeOH, (ii) TFA, DCM, (iii) [IrCl <sub>2</sub> Cp*] <sub>2</sub> , Et <sub>3</sub> N, MeOH. ....	116
<b>Scheme 3:</b> The 3 steps synthesis of ketone <b>30</b> . ....	117
<b>Scheme 4:</b> Claisen condensation scheme of ketone <b>30</b> and diethyl oxalate, for the synthesis of product <b>31</b> . ....	118
<b>Scheme 5:</b> Representation of a retro-Claisen reaction and a self-Claisen condensation mechanism, adapted from the proposed mechanism by Ji and co-workers. <sup>233</sup> .....	120
<b>Scheme 6:</b> NaOEt catalysed Claisen condensation between ketone <b>30</b> and diethyl oxalate. Representation of the by-products created by self-Claisen condensation and retro-Claisen reaction, based on Ji and co-workers proposed mechanism ( <b>Scheme 5</b> ). <sup>233</sup> .....	121
<b>Scheme 7:</b> Example of the stabilisation mechanism of <i>t</i> -BuOLi mediated Claisen condensation, adapted from suggestions by Ji and co-workers. <sup>233</sup> .....	121
<b>Scheme 8:</b> Adapted protocol for the synthesis of product <b>31</b> by Claisen condensation. ....	122
<b>Scheme 9:</b> TFA catalysed deprotection of Boc intermediates <b>28</b> and <b>30</b> to afford the deprotected carboxylic acid <b>36</b> and ketone <b>37</b> . ....	124
<b>Scheme 10:</b> The two steps synthetic pathway to produce the ethyl4-(3-(N-(2-(( <i>tert</i> -butoxycarbonyl) amino)ethyl)sulfamoyl)phenyl)-3-methyl-2,4-dioxobutanoate <b>40</b> , following previously describe protocols. ....	126
<b>Scheme 11:</b> Hypothesis of a degradation mechanism of compound <b>31</b> into the deprotected ketone <b>37</b> , following a decarbonylation-decarboxylation combination. ....	127

<b>Scheme 12:</b> Ester deprotection of product <b>31</b> through saponification reaction to afford the product <b>38</b> . The hydrophilic anchoring part was composed of ketones and carboxylic acid.....	128
<b>Scheme 13:</b> Proposed retro-Claisen condensation mechanism during the ester deprotection reaction. The high reactivity of the second carbonyl resulted in a nucleophilic attack by hydroxyl group, promoting a C-C bond cleavage and the production of an enolate, subsequently reprotonated to generate the ketone <b>30</b> . .....	129
<b>Scheme 14:</b> Synthetic pathway using Cbz protection. The deprotected ketone <b>37</b> was first re-protected by Cbz group before the synthesis of the $\beta$ -diketoester <b>42</b> . The final step was the Cbz deprotection catalysed by Pd/C.....	130
<b>Scheme 15:</b> Proposed mechanism for the hydrogenolysis by H <sub>2</sub> of the $\beta$ -diketoester anchoring part, catalysed by Pd/C. <sup>263</sup> .....	131
<b>Scheme 16:</b> Example of observed by-products obtained from the Cbz deprotection of $\beta$ -diketoester <b>42</b> using Pd/C - H <sub>2</sub> hydrogenation. ....	131
<b>Scheme 17:</b> The LAH-mediated reduction of compound <b>31</b> afforded the Boc-protected keto-diol <b>43</b> . Both aliphatic ketone and final ester were reduced to hydroxyl groups.....	132
<b>Scheme 18:</b> TFA-mediated <i>N</i> -Boc deprotection of the ketodiol <b>43</b> , using the previously described protocol. The resulting product <b>44</b> was a deprotected furan. ....	134
<b>Scheme 19:</b> Top: generally accepted Paal-Knorr mechanism for the acidic catalysed cyclisation of diketone into furan. Bottom: Reported Paal-Knorr pathway by Amarnath and Amanarth. <sup>276,277</sup> .....	136
<b>Scheme 20:</b> Proposed mechanism of the TFA-catalysed furan cyclisation to form the furan product <b>44</b> . The mechanism was based on reported Paal-Knorr pathways. <sup>276,277</sup> .....	136
<b>Scheme 21:</b> The LAH-mediated reduction of the Cbz-protected product <b>42</b> afforded the Cbz-protected keto-diol <b>45</b> . Both aliphatic ketone and final ester were reduced to hydroxyl groups.....	137
<b>Scheme 22:</b> Synthesis of the triol <b>46</b> by Cbz deprotection of product <b>45</b> , using a reductive cleavage by hydrogen gas catalysed by palladium on carbon.....	138

**Scheme 23:** Transfer hydrogenation of 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline to produce the chiral amines 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroquinoline, catalysed by the artificial metalloenzymes or the iridium catalysts alone, adapted from a previously describe protocol by Ward and co-workers.<sup>290</sup> ..... 180

## List of Abbreviations

ADH(s)	Alcohol Dehydrogenase(s)
AM(s)	Artificial metalloenzyme(s)
Boc	<i>tert</i> -Butyloxycarbonyl
BSA	Bovine serum albumin
Cbz	Benzyloxycarbonyl
Cp*	Pentamethylcyclopentadiene
COSY	Homonuclear correlation spectroscopy
DCM	Dichloromethane
DFT	Density functional theory
DMSO-d6	Deuterated dimethyl sulfoxide
<i>E. coli</i>	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
ER	Enoate reductase
ESI-MS	Electrospray ionisation mass spectrometry
EtOAc	Ethyl acetate
FEP	Free energy perturbation
His-tag	Hexahistidine tag
HLADH	Horse Liver Alcohol dehydrogenase
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear single quantum correlation
IC <sub>50</sub>	Inhibition concentration 50 %
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K <sub>M</sub>	Michaelis constant
LB	Lysogeny broth
LDR	Long chain reductase
m / z	Mass to charge ratio
MAO	Monoamineoxidase
MD	Molecular dynamics

MDR	Medium chain reductase
MeOH	Methanol
MM	Molecular mechanics
MWCO	Molecular weight cut-off
NAD <sup>+</sup> / NADH	Nicotinamide adenine dinucleotide oxidised form / 1,4 reduced form
NADP <sup>+</sup> / NADPH	Nicotinamide adenine dinucleotide phosphate oxidised form / 1,4-reduced form
NHC	<i>N</i> -heterocyclic carbene
NMR	Nuclear magnetic resonance
OD (OD <sub>600</sub> )	Optical density (measured at 600 nm)
OYE	Old yellow enzyme family
PDB	Protein Data Base
PBS	Phosphate Buffered Saline
ppm	Parts per million
QM	Quantum mechanics
RMSD	Root Mean Square Deviation
rpm	Rotation per minute
SDR	Short chain reductase
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TbADH	<i>Thermoanaerobacter brokii</i> Alcohol Dehydrogenase
TFA	Trifluoroacetic acid
TON	Turnover number
Tris.HCl	Tris(hydroxymethyl)aminomethane chloride
U / mg	Specific unit of activity = $\mu\text{mol}$ substrate per minute per mg of enzyme
UV-Vis	Ultraviolet and visible light
v / v	Volume / volume ratio
VdW	Van der Waals
V <sub>max</sub>	Maximum reaction rate (velocity) of enzyme
WT	Wild type

# 1 Literature review

## 1.1 Introduction

Challenges in the production of high-value compounds for the chemical and pharmaceutical industry includes the development of green processes. The increasing production of toxic and hazardous wastes needs actions for the protection of humans and the ecosphere. More sustainable and environmentally friendly techniques are now required, alongside a low cost and efficient synthesis.

Biocatalysis stands as an appealing technology and a key solution for the coming years. The use of biological entities, such as enzymes, reduces the application of chemical products (e.g. derived from fossil fuels), while bringing selectivity and decreasing the amount of waste (e.g., biodegradable catalyst, no chemical reagents, no side reactions).

Biocatalysis is an attractive and expanding field also due to the necessity to produce enantiopure products. Based on improvements in protein engineering technologies (e.g. bioinformatics, genome sequencing, directed evolution, chemical optimisation etc.), biocatalysis is in constant innovation, with an evolution in the portfolio of available biocatalysts.

Nowadays, the increasing use of computational aided approaches has improved the protein engineering capacities further.<sup>1,2</sup> Numerous studies aimed to construct advanced biocatalysts by re-designing protein active sites, or by *de novo* construction of catalytic sites.<sup>3</sup> The focus of these studies is the creation of biocatalysts with novel, non-natural activities.

A particularly promising strategy is the introduction of a non-native metal ion inside a protein scaffold through various anchoring methods. The resulting product is called an artificial metalloenzymes (AMs). These biocatalysts combine the non-natural catalytic activity of a given transition metal catalyst with the controlled environment of natural enzymes.<sup>4</sup> These entities are able to address challenges of green chemistry: reaction control, renewable catalyst, enantioselectivity and environmentally friendly conditions. Studies already proved the construction of efficient AMs.<sup>5</sup> However, only a few attempted the design of AMs by computational methods. On the other hand,

the drug design literature offers many examples where non-natural chemical structures are docked within proteins using supramolecular interaction for inhibition purposes.<sup>6</sup> This strategy could be investigated to build AMs accommodating metal catalysts in the binding site.<sup>7</sup>

In the industrial biocatalysis field, a better design of AMs would help the development of a truly adaptable catalyst, addressing some of the green chemistry challenge and enhancing the selective synthesis of compounds.

Hence, this project proposes the development of an AMs catalysing the transfer hydrogenation of imines for the synthesis of chiral amines. This functional group is commonly found in natural and synthetic products, where it can be responsible for biological activity. Chiral amines are of importance in the pharmaceutical industry, where they are used as building blocks in the synthesis of many drugs. The production of chiral amines is generally achieved through asymmetric synthesis. Therefore, the use of AMs for the creation of chiral amines is a good opportunity to develop a safer process that fulfils the requirements of green chemistry.

## 1.2 Artificial metalloenzymes: a strategy for biocatalysis

### 1.2.1 *Background*

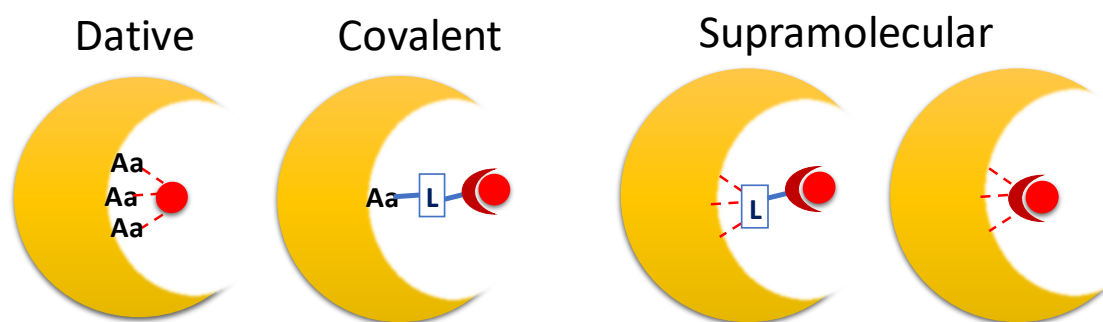
In biocatalysis domain, one attractive strategy to achieve non-natural activities in biological environment is the use of AMs. These advanced hybrids are created by the incorporation of a metal ion in a protein or enzyme scaffold. AMs make use of the enzyme scaffold selectivity and the reactivity of a non-native metal catalyst. By variation of the metal complex nature, chemists can adapt the AMs to their targeted reaction. Thus, AMs support a larger range of biological and non-natural reactions in aqueous media, while providing a better stereoselectivity recognition of substrate and products.<sup>4</sup> Several studies have already proven this concept and a variety of enantioselective catalysis reactions have emerged.<sup>8</sup> One prominent example of efficient AMs is the biotin-streptavidin technology developed by the Ward group, which led to the development of artificial asymmetric transfer hydrogenases with high stereoselectivity.<sup>9</sup>

Three important parameters describe the unique nature of AMs: the host (enzyme structure), the transition metal complex and the interaction between these two entities. The transition metal represents the first coordination sphere, responsible for the catalytic activity. The metal must be inserted orthogonally inside the enzyme structure to avoid interferences with the selective functionalities of the protein scaffold. The combination of interactions between hosts and the first coordination sphere is known as the second coordination sphere. This is a key parameter in the stability of the complex and the efficiency of the AMs.<sup>10</sup> Careful design and strong interactions form a stable enzyme-metal catalyst complex, and thus a robust AM in its reactivity and selectivity. The evaluation of space inside the enzyme scaffold and the mode of anchoring of the metal catalyst are therefore part of prerequisites parameters necessary to engineer an efficient AM.<sup>8,11</sup>

The engineering of AMs is challenging in the design of the two coordination spheres. For the 1<sup>st</sup> coordination sphere, evaluation of space inside the enzyme is a prerequisite at the introduction of a reactive transition metal catalyst able to fit. In the 2<sup>nd</sup> sphere, the mode of anchoring, the various interactions for a stable binding or the proximity with the substrate are important choices.

### 1.2.2 *Metal anchoring strategies*

The introduction of a metal or a metal complex inside the protein scaffold adopts different anchoring approaches (**Figure 1**): dative, covalent or supramolecular anchoring.<sup>4</sup>



**Figure 1:** Representation of metal anchoring methods, to build AMs. The following codes were applied: the amino acid (Aa), the covalent or supramolecular ligands and spacers (L), the ligand coordinating the metal ion (red crescent), the metal ion catalyst (red ball). The dashed lines represent non-covalent interactions, and the plain lines represent covalent interactions.



#### 1.2.2.1 Dative anchoring

In the dative anchoring, a binding site is built inside the natural enzyme structure to accommodate the transition metal catalysts (**Figure 1**). Several amino acids are known to establish dative bonds with transition metal catalyst through N, O and S coordination. Thus, the protein scaffold can be arranged into an efficient three-dimensional binding site. Two techniques have been used: the repurpose of a natural enzyme metal binding site or the creation of a new catalytic binding site.<sup>12</sup>

##### - *Repurpose of natural metalloenzymes*

Almost a third of enzymes in nature are metalloenzymes.<sup>13</sup> They have naturally evolved to hold the characteristics for an efficient metal binding in their structures (e.g. Zn-dependent alcohol dehydrogenases, Fe-containing cytochromes, Cu-laccase...). Thus, natural metalloenzymes are perfect candidates for the substitution of their native metal by a transition metal catalyst.

The technique consists in removing the natural metal from the enzyme (i.e. dialysis) and mixing the new metal ion with the apoprotein. In 1976, Kaiser and Yamamura first reported this technique using a Cu(II) ion to substitute the zinc of a Zn(II)-dependent carboxypeptidase A.<sup>14</sup> The newly engineered AM was able to catalyse the oxidation of ascorbic acid. Several other studies involved different enzyme scaffolds, including a Zn(II)-carbonic anhydrase with Mn(II) replacing the native metal,<sup>15,16</sup> Pd<sup>2+</sup> ions inserted inside an apo-ferritin<sup>17,18</sup> or two Cu<sup>2+</sup> embedded in a repurposed di-Zn metallo- $\beta$ -lactamase.<sup>19</sup> All these enzymes showed new to nature catalytic reactivity.

This strategy is however limited in the scope of the binding site where only metalloenzymes with a proper coordination motif for a metal can be used. In addition, the enzymes need to be separated from their natural metal with certainty. If the metal is not removed, the new transition metal cannot be inserted. Finally, metal substitution enhances the non-specific metal binding to other potential binding sites.

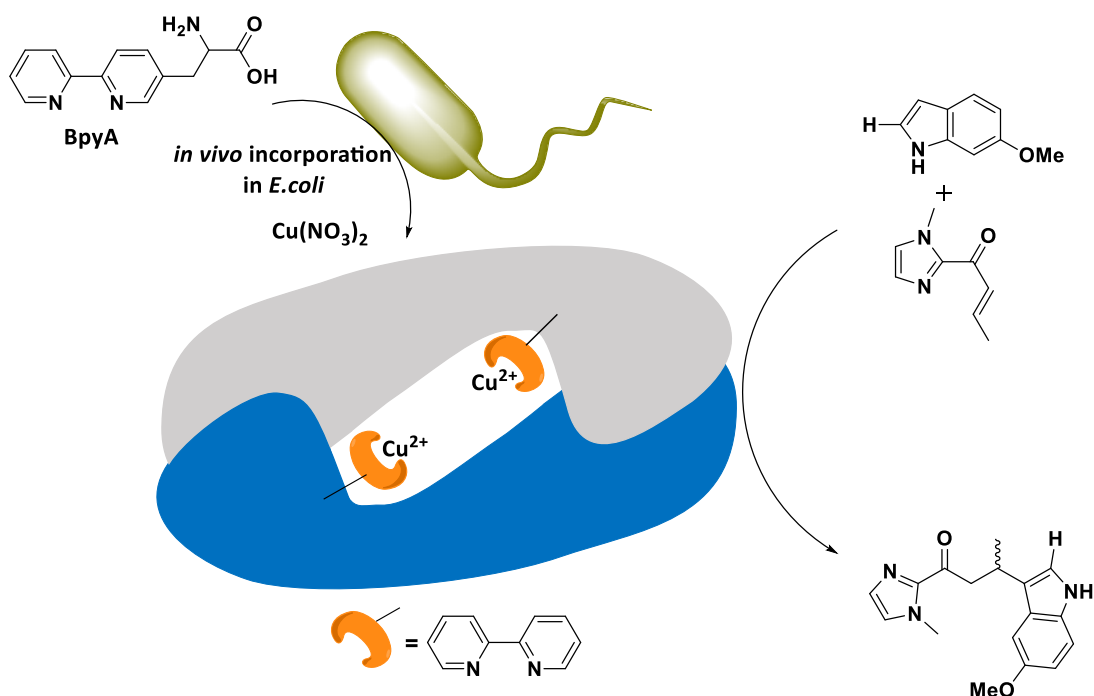
##### - *Introduction of a new metal-binding site*

One strategy to overcome the limitation of scaffold scope is the screening of existing protein binding sites for a potential binding of transition metals.<sup>12</sup> Ward and co-

workers used this technique to create a new to nature OsO<sub>4</sub>-streptavidin AM, after the screening of five potential proteins.<sup>20</sup> The second option for the creation of dative anchoring sites is the *de novo* design of a metal binding protein. A new metal-binding site is formed by the incorporation of the appropriate amino acids (or unnatural amino acids) *via* site-directed mutagenesis. This strategy significantly increases the choice of scaffolds.

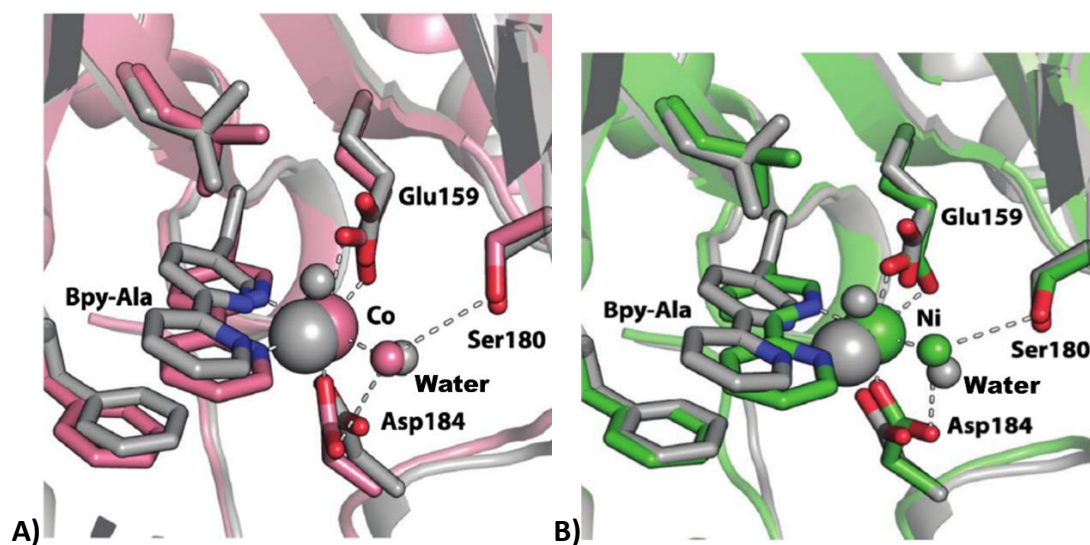
An example of direct metal incorporation was provided by Reetz and co-workers, who created a Cu(II)-binding site inside the thermostable enzyme tHisF.<sup>21</sup> They first inspected the protein crystal structure to target amino acids for the conventional site-specific mutagenesis. The authors introduced a His/His/Asp coordinating environment in the TIM-barrel of the tHisF to welcome the Cu(II) ion. The resulting AM was able to perform a new Diels-Alder reaction with moderate enantioselectivity (46 % ee).

To expand the opportunities of metal coordination in scaffolds, researchers have also developed unnatural amino acids. As an example, Roelfes and co-worker inserted *in vivo* the unnatural amino acid bipyridyl alanine (BpyAla) inside their extensively studied AM scaffold, the lactococcal multidrug resistant receptor (LmrR, **Figure 2**).<sup>22</sup> This dimer receptor forms a hydrophobic pocket that fits both the metal complex and the substrate. Insertion of BpyAla coordinates the binding of Cu(II) ion to perform a Friedel-Craft reaction catalysis with up to 94 % ee.



**Figure 2:** Representation of the LmrR engineered *in vivo* with incorporation of BpyAla residues to create AM. After coordination of the  $\text{Cu}(\text{II})$  ion, an example of its new to nature catalysis of Friedel-Crafts alkylation reaction is presented. Adapted from Roefle article.<sup>23</sup>

With help from computational design RosettaMatch,<sup>24</sup> Baker and co-workers also created a new AM using BpyAla, aspartic acid and glutamic acid residues as metal-binding amino acids (**Figure 3**).<sup>25</sup> Experimental confirmation revealed a crystal structure of the designed protein, with the BpyAla binding  $\text{Co}^{2+}$  or  $\text{Ni}^{2+}$ .



**Figure 3:** Reproduction of Baker and co-workers designed protein.<sup>25</sup> **A)** Solved crystal structure of their designed protein with  $\text{Co}^{2+}$  **B)** and  $\text{Ni}^{2+}$ .

The development of methodologies for the *de novo* binding site design has been extensively studied to increase the potential of AMs engineering. However, the

engineering of a new binding site must not prefer the metal binding over other structures the scaffold needs to accommodate, like the substrates and any potential cofactors. In addition, accommodation of the new binding site can be a challenge for the overall structural stability.

#### 1.2.2.2 Covalent anchoring

In this second approach, a synthetic ligand linked to a metal catalyst is covalently bound to an exposed amino acid inside a protein scaffold (**Figure 1**). This results in the most stable and the most common linkage between two structures. Besides, this strategy provides high flexibility in the scaffold and transition metal choice.

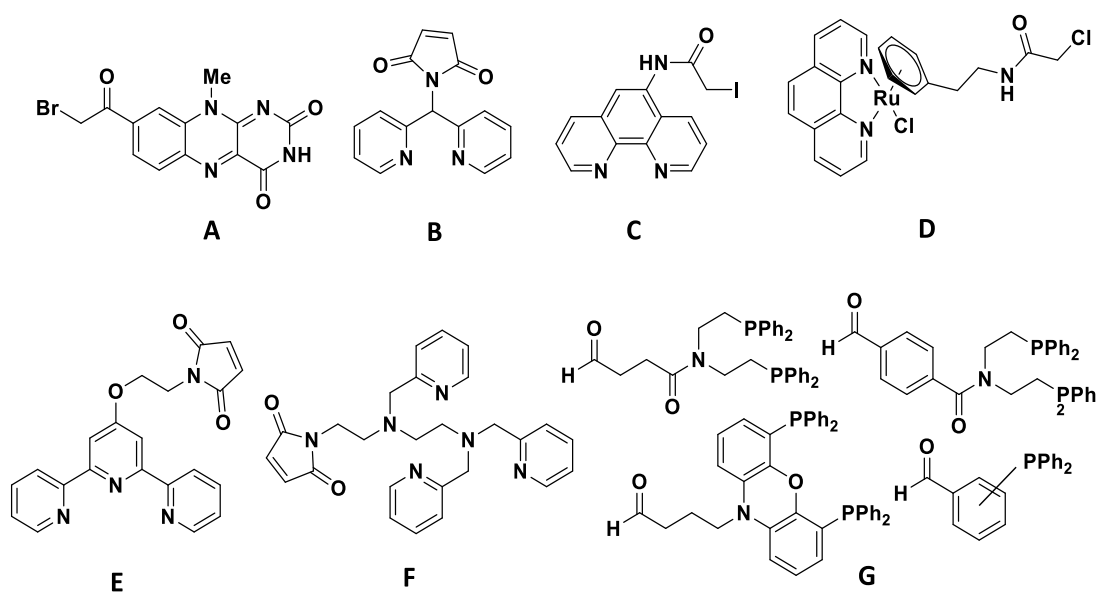
This bioconjugation still implies requirements. First, a suitable reactive residue (generally nucleophile cysteine or lysine) must be present in the binding site while similar reactive residues have to be removed. Secondly, the ligand should possess a reactive functional group able to make the irreversible connection reaction (e.g. nucleophilic attack, disulphide bond, cycloaddition). Finally, the enzyme cavity must be large enough to contain the whole metal complex with the ligand linker and the substrate.<sup>11</sup>

Kaiser and co-worker first reported the potential of a synthetic ligands covalently bind to a protein scaffold.<sup>26</sup> They used the nucleophilic activity of the single cysteine in papain to selectively alkylate a brominated flavin  $\alpha$ -haloketone (**Figure 4, A**). This resulted in an oxidoreductase catalysing the oxidation of dihydronicotinamides.

Distefano and co-workers first started to explore enantioselective catalysis with covalent binding of a metal catalyst. They focused on the adipocyte lipid-binding protein which structure contains only one cysteine.<sup>27</sup> They bioconjugated the iodoacetamido-1,10-phenanthroline ligand and metalated the bioconjugate with Cu(II) (**Figure 4, B**). The resulting AM catalysed amide and ester hydrolysis with high enantioselectivity (more than 90 % ee).<sup>28</sup>

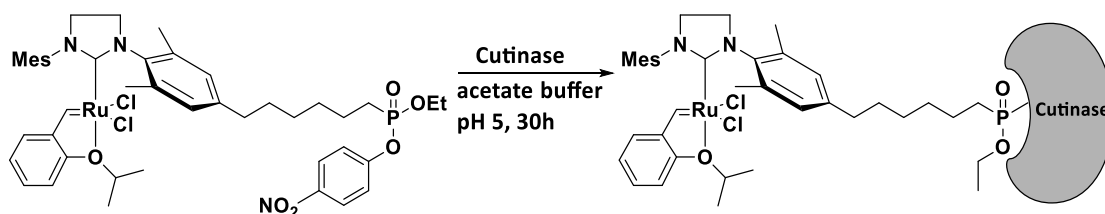
Based on these preliminary results, several groups have tested multiple ligands and protein scaffolds for covalent bioconjugation. Reetz and co-workers developed a series of maleimide-substitutes bipyridine ligand metalated with Pd, Cu and Rh (**Figure 4, B and C**). They proved the alkylation with first the papain cysteine<sup>29</sup> and

then inside the tHisF thermostable enzyme.<sup>30</sup> Salmain and co-workers then bioconjugated a Ru(II) $\eta^6$ -arene complex inside papain protein with significantly better catalytic activity on Diels-Alder reaction than preliminary work from Kaiser and co-workers (**Figure 4, D**).<sup>31</sup> Lewis and co-workers<sup>32</sup> and Hayachi and co-workers<sup>33</sup> designed a maleimide-substituted terpyridine ligand metalated with Mn and Cu respectively (**Figure 4, E**). They both bioconjugated inside the nitrobindin protein to produce AMs catalysing oxygenations and Diels-Alder reactions. Meanwhile, Banse and co-worker bioconjugated a Fe(II) amine/pyridine ligand complex with the only cysteine available inside the bovine  $\beta$ -lactoglobulin (**Figure 4, F**), resulting in a peroxygenase.<sup>34</sup>



**Figure 4:** Representation of ligands structures used in literature to make covalent linkage in order to anchor metal ions inside a host to form AMs.

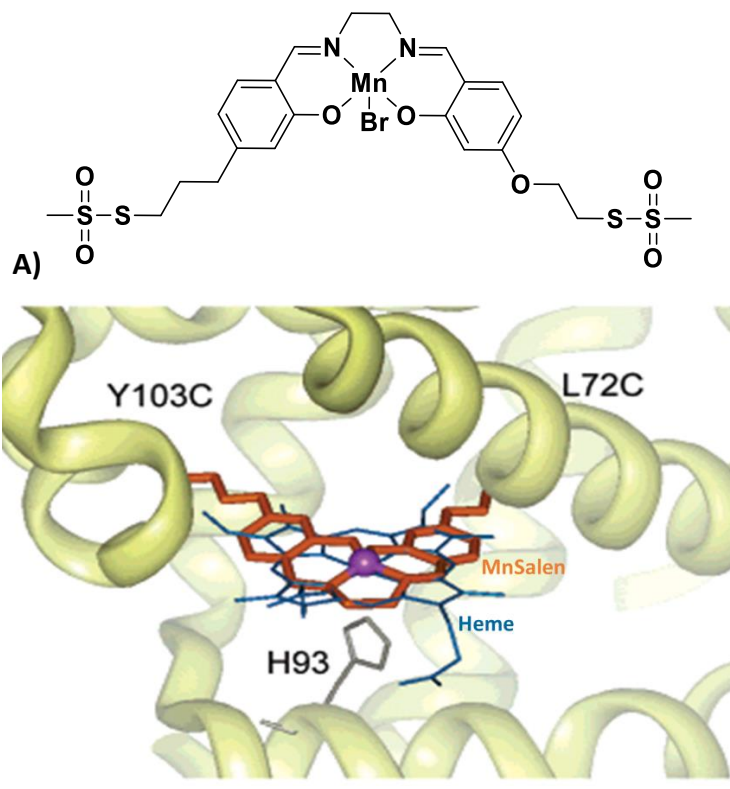
Some non-maleimide based ligands have also been reported. For example, Kamer and co-worker bioconjugated a phosphine-based ligand, metalated with Rh or Pd, inside the photoactive yellow protein (**Figure 4, G**).<sup>35,36</sup> Gebbink and co-workers reported the bioconjugation inside a lipase cutinase of a Grubbs type metal catalyst: Rh(NHC) or Ru(NHC) (NHC= N-heterocyclic carbene).<sup>37,38</sup> The authors used a phosphonate lipase inhibitor as ligand to covalently bind the Grubbs catalyst through active site-directed hybridization of a serine residue in the lipase (**Figure 5**). The resulting AMs catalysed the hydrogenation of olefin and ketone, the ring-closing metathesis of diallyl tosylamine and the cross-metathesis of allylbenzene.



**Figure 5:** Reaction scheme of the bioconjugation of lipase cutinase hybrid, adapted from Gebbink and co-worker.<sup>37</sup>

These authors also realised that the chain length of their ligand was important. With a short linker, no metathesis activity was spotted.

However, since Distefano works, little or no enantioselectivity has been detected with all engineered AMs. The importance of the ligand size and binding position is essential for selectivity. Lu and co-worker followed a rational design approach on the anchoring location and movements of a covalently bound ligand.<sup>39</sup> They studied a MnSalen complex bioconjugated inside an apo-myoglobin (Mb) (**Figure 6, B**).<sup>40</sup> The authors predicted, by computational modelling, several ligand location prior to mutating the targeted residues inside the heme cavity.



**Figure 6:** MnSalen complex inside apo-Mb protein scaffold, reproduced from Lu and co-workers.<sup>40</sup> **A)** MnSalen complex **B)** Computer model of MnSalen dual anchoring in Mb, overlayed with the protein natural heme cofactor.

The resulting optimised AMs carried a dual anchoring to two mutated cysteine residues. The experimental work showed a clear increase in asymmetric sulfoxidation (up to 83 % ee). In addition, any changes in anchoring positions revealed significant changes in the reaction rate and enantioselectivity. On sulfoxidation of thioanisole, the MnSalen complex alone yielded 1 % ee-(*S*), a single anchoring yielded 12 % ee-(*S*) and for the dual anchoring 66 % ee-(*S*) was recorded. The two-point covalent anchoring may have decreased the complex movement within the scaffold, enhancing selectivity.

This hypothesis was later confirmed by Lewis and co-workers,<sup>41</sup> who introduced a histidine residue to coordinate their metal catalyst in a two-point anchoring. The authors extended activity and selectivity of propyl oligopeptidase AMs by an active site mutagenesis. Roelfes and co-workers clearly confirmed the impact of the anchoring position on the conversion and enantioselectivity by a mutagenesis study of their Cu(II)-LmrR AM.<sup>42,43</sup> The environment around the metal complex catalyst, provided by the protein scaffold, is key to an efficient catalysis.<sup>23,43</sup>

Overall, the covalent anchoring of a metal catalyst introduced non-natural reactions, increasing reaction scope of enzymes, while adding high enantioselectivity and robustness of the complex metal-enzyme. However, the covalent anchoring of a metal complex presents challenges. For example, a single residue is needed for the bioconjugation. Prior steps are thus necessary to mutate any potential other reactive amino acid in the host scaffold, with caution for the structural stability. Besides, the chemical reaction endured by the enzyme during bioconjugation also threaten the integrity of the scaffold. Finally, the ligand size may limit the encapsulation inside the protein scaffold.

#### 1.2.2.3 Supramolecular anchoring

Covalent and dative strategies are limited by the need of specific amino acids and restricting chemical reactions. Another strategy for the incorporation of a metal complex catalyst is the supramolecular anchoring (**Figure 1**). In this approach, the metal catalyst is anchored by strong and specific non-covalent linkage inside the protein. These types of linkages involve no electron sharing, but weaker interaction forces such as hydrophobic, electrostatic interactions or hydrogen bonding.

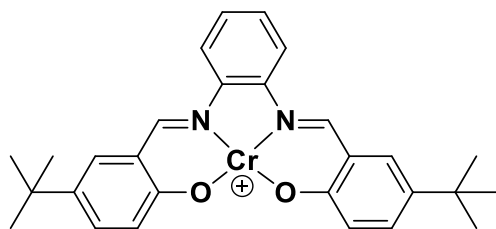
This supramolecular strategy has several advantages: there is no need for a coupling chemical reaction, no need for modifications of the enzyme scaffold, diversity in the nature of the metal catalyst and enzyme structure and easier optimisation of the engineered AMs.<sup>44</sup> There are however some drawbacks in the design of supramolecular AMs. For example, the strength and selectivity of the non-covalent bindings are a challenge, with possible formation of unexpected bindings inhibiting the enzyme activity.

The supramolecular binding approach can be divided in two strategies: the metal complex is directly bound to the protein scaffold or a ligand linker bind the metal complex to the scaffold.

- ***Direct metal complex supramolecular binding***

Unlike metal binding using dative interactions, the metal complex binding involves supramolecular interactions with a ligand coordinating the metal ion. These interactions are often labile, and the slightest modifications in the ligand structure results in an entire change of reactivity.

The apo version of heme proteins were among the first to be shown to accommodate non-natural metal complexes, which were non-covalently bound in place of heme. Watanabe and co-workers introduced several metal complexes inside a suitable mutant of apo-Mb protein.<sup>45</sup> They generated artificial sulfoxidases by the introduction of a Mn(III)/Cr(III)-salophen Schiff bases (**Figure 7**). The planar structure of the metal complex results in similar binding as the natural cofactor heme.



**Figure 7:** Example of Cr(III)-salophen catalyst complex, reported from Watanabe and co-workers.<sup>45</sup>

Hartwigs and co-workers created several AMs using protoporphyrin and mesoporphyrin IX (PIX) as metal coordinating ligand. They reconstituted apo-(PIX)-proteins by replacing the natural iron metal. In the apo-Mb scaffold, they inserted a protoporphyrin IX metalated with different metals (Co, Cu, Mn, Rh, Ir, Ru and Ag).<sup>46</sup> After optimisation by targeted mutations of the active site, the best AM selected was

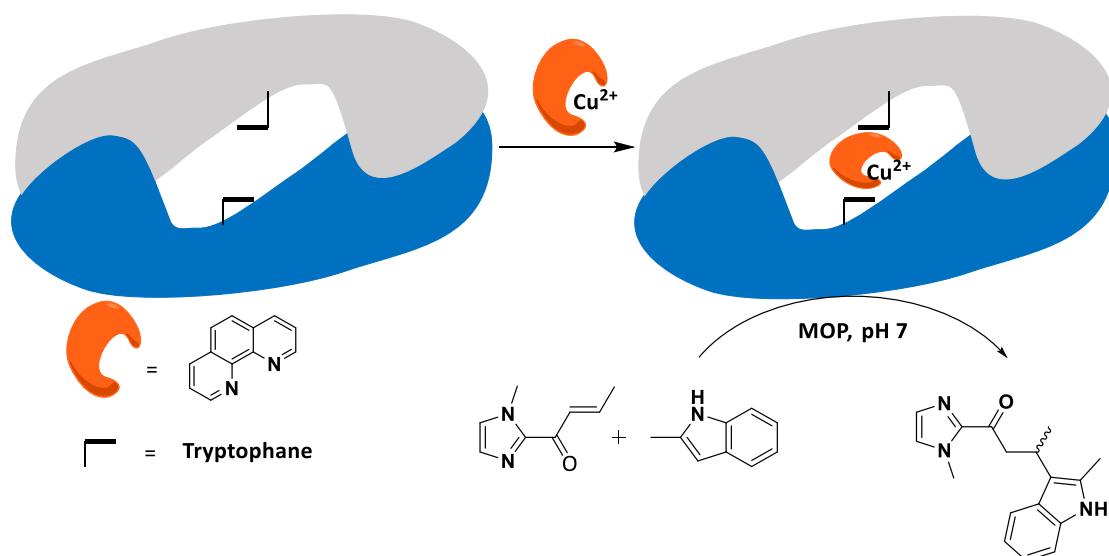


the Ir(Me)-PIX-Myo, catalysing the epoxidation of olefin (up to 82 % ee) and C-H intramolecular insertion of diazo substrate (up to 86 % ee). Based on this work, Hartwig and co-workers recently replaced a cytochrome P450 heme by the Ir(Me)-PIX catalyst, using molecular evolution.<sup>47</sup> The new AM showed non-natural reactivity of intramolecular C-H amination of sulfonyl azides (**Figure 8**).



**Figure 8:** C-H amination reaction scheme catalysed by the best Ir(Me)-PIX CYP119 enzyme variant. Figure adapted from Hartwig and co-workers.<sup>47</sup>

Another scaffold for metal complex binding is the LmrR, extensively studied by Roelfes and co-worker. The authors recently reported the binding of Cu(II)-phenantroline catalyst inside the interface cavity of the receptor (**Figure 9**).<sup>48</sup> The hydrophobicity of the cavity and two central tryptophans favour a planar coordination of the metal complexes. They thus started to evaluate the activity of several bidentate aromatic nitrogen Cu(II) complexes. The engineered AMs were able to catalyse a Friedel-Crafts alkylation of indoles with high enantioselectivity (up to 93 % ee).



**Figure 9:** Representation of supramolecular assembly of Cu(II)-phenantroline complex catalyst inside LmrR, with an example of its new to nature Friedel-Crafts alkylation of indole. Adapted from Roelfes and co-workers.<sup>23</sup>

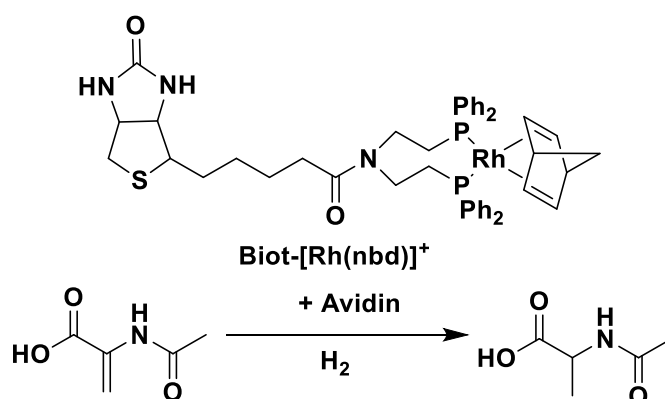
The Roelfes group later bonded a hemin complex inside LmrR for an AM catalysing cyclopropanation reactions with enantioselectivity (51 % ee).<sup>23</sup>

- **Ligand linkers mediate supramolecular metal complex binding**

In the previous non-covalent metal complex binding strategy, modifications to the ligand coordinating the metal will automatically change the binding to the scaffold and thus the activity and selectivity of the AM. In turn, the binding of the metal complex might affect its chemical abilities.

One way to decouple the catalyst from the binding process is to add an anchoring ligand in between. This approach also brings challenges: the ligand must bind strongly enough to the scaffold and the metal position must allow reactivity and selectivity.

One of the first and the most studied scaffold to generate non-covalent AM is the avidin (Av) or streptavidin (Sav). This approach relies on the naturally strong non-covalent interaction between the protein (strept)avidin and its natural coenzyme biotin ( $10^{12}$  -  $10^{15}$  M<sup>-1</sup> affinity).<sup>49</sup> The concept was first described by Wilson and Whitesides in 1978.<sup>50</sup> They embedded the achiral [Rh(nbd)(diphosphine)]<sup>+</sup> complex (nbd= norbomadiene) within avidin to generate an artificial hydrogenase (**Figure 10**). To this end, the authors substituted the biotin anchoring ligand with a diphosphine to coordinate Rh(I). The resulting AM catalysed the reduction of  $\alpha$ -acetamidoacrylic acid with enantioselectivity 41 % ee-(S).

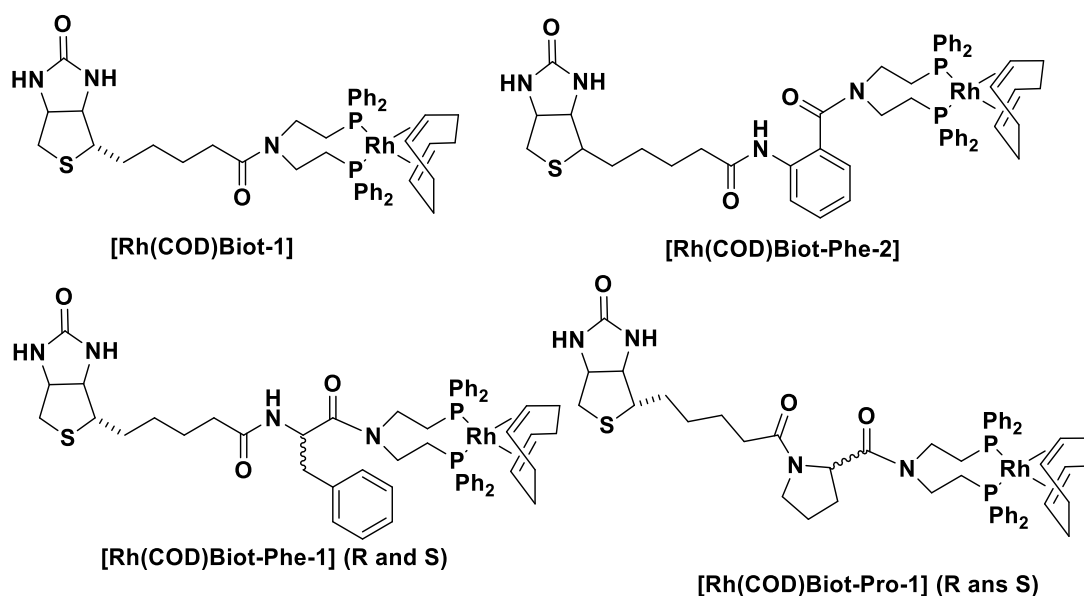


**Figure 10:** Presentation of the first AM based on Av/Sav technology, catalysing the hydrogenation of  $\alpha$ -acetamidoacrylic acid. This AM was composed of a biotinylated Rh catalyst non-covalently bound inside avidin. Figure adapted from Wilson and Whitesides.<sup>50</sup>

Later, Chan and co-worker reported similar conclusions using chiral biothiynylated Rh(I)-pyrphos complex linked to biotin and embedded within avidin.<sup>51</sup> The resulting AM catalysed the asymmetric hydrogenation of itaconic acid (37 % ee-(S)).

In recent years, the biotin-(strept)avidin technology has been extensively developed by Ward and co-workers. They started in 2001 by changing the scaffold of the protein to Streptavidin (Sav), easily recombinantly expressed in *E. coli*. The authors then applied combined optimisations. A chemical optimisation of the synthetic ligand gave a small library of metal catalyst complexes.<sup>49</sup> Besides, genetic modifications created a library of Sav mutants, changing the surrounding of the metal catalyst.<sup>52</sup> Combination of both first and second coordination spheres optimisations engineered a small library of efficient AMs.

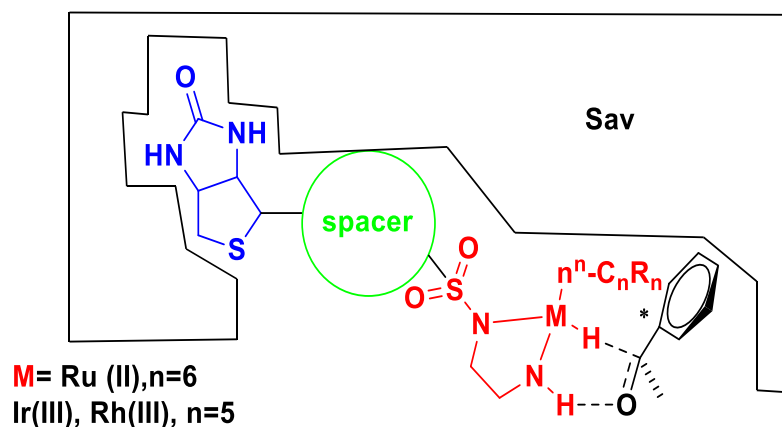
Ward and co-workers initial experiments bound biotinylated Rh-diphosphine complexes ([Rh(COD)Biot], COD= 1,5-cyclooctadiene) within Sav, to build an artificial hydrogenase catalysing the hydrogenation of N-protected dehydroamino acids. They then chemically optimised the first-generation catalyst by adding amino acid linkers between biotin and the metal complex (18 biotinylated catalysts designed, **Figure 11**). They also realised saturation mutagenesis for the creation of a Sav mutant library. The combined strategy improved the catalytic selectivity up to 96 % ee-(*R*) and 95 % ee-(*S*) with [Rh(COD)Biot-Phe-2]-Sav(S122G).<sup>53,54,55</sup>



**Figure 11:** Examples of biotinylated metal complexes, designed to be non-covalently embedded inside Sav to build artificial hydrogenases. [Rh(COD)Biot-1] is constructed with the first generation of diphosphine ligand. The other complexes were then optimised with amino acid spacers.<sup>49,55</sup>

Following their success in hydrogenation catalysis, Ward and co-workers applied the biotin-streptavidin technology to one fundamental chemical reaction: the asymmetric transfer hydrogenation of carbonyl. In addition, the authors investigated

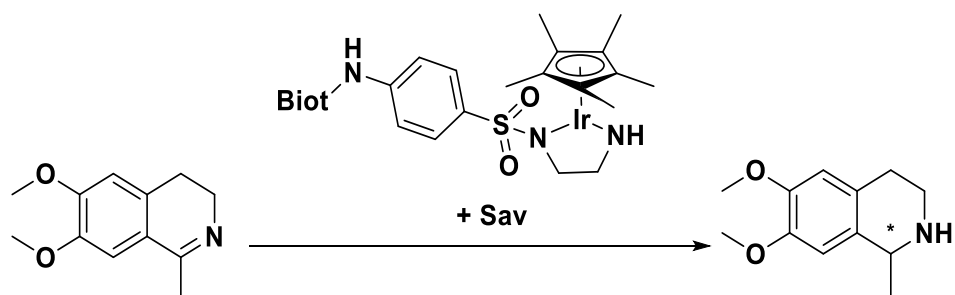
the versatility of their AMs to imitate the ketone oxidoreduction of a natural enzyme. They selected the piano-stool catalyst complex, which was used by Noyori to catalyse the prochiral ketone reduction and the oxidation of secondary alcohol.<sup>56</sup> Ward and co-workers embedded diamine biotinylated  $d^6$  piano-stool complexes inside Sav (**Figure 12**).<sup>57</sup> They again combined chemo-genetic optimisations to create a library of improved AMs.<sup>58,59</sup>



**Figure 12:** AM model for the enantioselective transfer hydrogenation of carbonyl, adapted from Ward and co-workers.<sup>58</sup> The ligand biotin (blue) is separated from the  $d^6$  piano-stool metal complex (red) by different spacers (green). The prochiral substrate receives two hydrogens without coordinating with the metal.

The common transition metals rhodium (Rh), iridium (Ir), and ruthenium (Ru) were used with different arene-cap. Variation in the capping impacted the enantiopreferences. In addition, distinct genetic mutations on the scaffold showed an impact on stereoselectivity when cationic residues were around the metal.<sup>52</sup> The best enantioselectivity was obtained with a  $[\eta^6\text{-}(p\text{-cymene})\text{RuCl-Biot}]$  - Sav S112Y mutant for 97 % ee-(*R*) and  $[\eta^6\text{-(benzene)}\text{RuCl-Biot}]$  Sav S112R mutant for 70 % ee-(*S*). For the reverse reaction, only the oxidation of sec-phenethyl alcohol was catalysed using the biotinylated amino-sulfonamide Ru(II) catalyst embedded within avidin protein.<sup>60</sup>

Ward and co-workers then utilised similar biotinylated  $d^6$  piano-stool catalysts for the creation of artificial cyclic imine reductases library (**Figure 13**). The chemo-genetic optimisations, coupled to computational design, enabled the selection of the best AMs with improved enantioselectivity: 85 % ee-(*R*) salsolidine with  $[\text{Cp}^*\text{IrCl-Biot}]$  - Sav S112R-N118P-K121A-S122M-L124Y and 78 % ee-(*S*) salsolidine with  $[\text{Cp}^*\text{IrCl-Biot}]$  - Sav S112A-N118P-K121A-S122M and an enhanced substrate scope.<sup>61,62,63,64</sup>

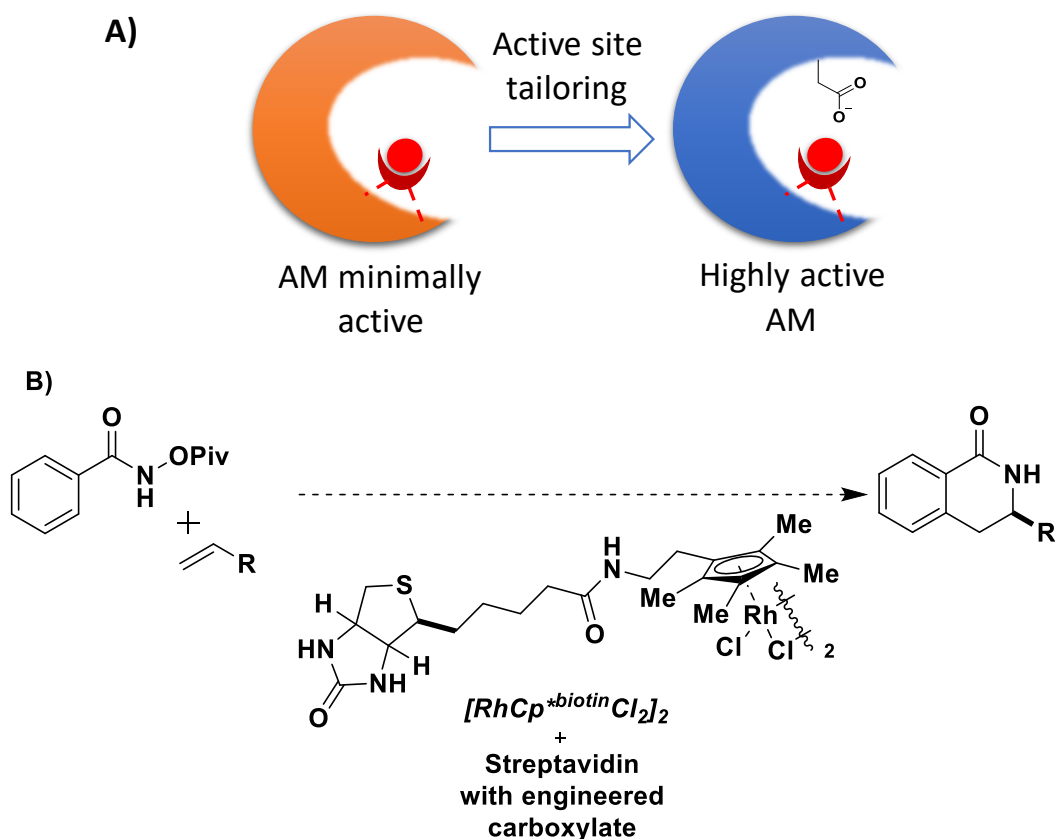


**Figure 13:** An example the asymmetrical reduction from cyclic imine to amines, catalyse by an AM based on the biotin-Sav technology.

Notably, Ward and co-workers used *in silico* studies to introduce a suitable histidine residue in the active site. The mutation provided a dual anchoring of the metal through a new dative bond, improving the activity and selectivity. The synthesis of this AM and the structural characterisation confirmed the computational design.<sup>65</sup>

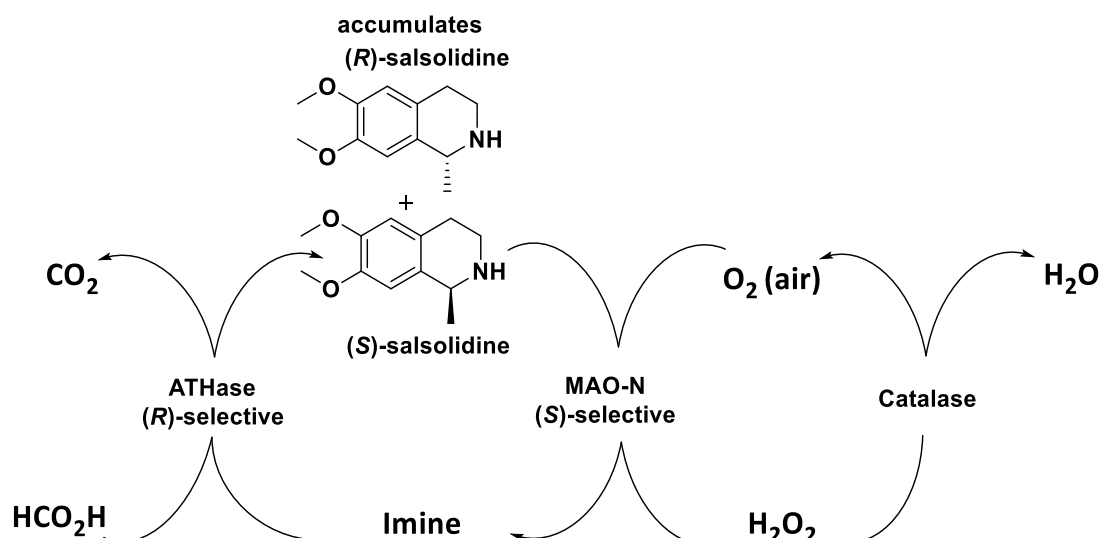
The authors applied the Biotin-Sav technology to numerous reactions, including notably olefin metathesis with Grubbs like catalysts,<sup>66</sup> sulfoxidation with Mn-salen catalysts<sup>67</sup> or phosphino-palladium catalysed Suzuki reactions.<sup>68</sup>

A collaboration with Rovis led to an AM catalysing the enantioselective benzannulation reaction (**Figure 14**). For a new C-C bond formation a C-H bond needs to be first activated by deprotonation from a base in excess. The authors induced an intramolecular deprotonation by introducing with site-directed mutagenesis a carboxylate residue in the active site.<sup>69</sup>



**Figure 14:** Enantioselective benzannulation catalysed by a new Sav-AM, adapted from Ward and co-workers.<sup>69</sup> **A)** Synergistic action of a basic residue introduced by site-directed mutagenesis with a biotinylated  $RhCp^*Cl$  catalyst non-covalently bound to Sav. **B)** Reaction scheme of the synthesis of dihydroisoquinolones by benzannulation reaction catalysed by the new Sav AM.

In nature, numerous enzymes are collaborating for the synthesis of products *via* metabolic pathways. However, if a non-natural occurring reaction is wanted in these cascade reactions, a free metal catalyst would not be able to stand alone. In their recent work, the Ward group hypothesised that AMs could be an attractive solution to protect metal catalysts during cascade reactions. To prove this, they first combined an artificial imine reductase with a natural amine oxidase (monoamine oxidase, MAO). However, the imine reductase was poisoned by hydrogen peroxide produced by the MAO. They thus introduced a third enzyme, a catalase (**Figure 15**).<sup>70</sup> Using a  $[Cp^*Ir(biot)Cl]$  – Sav S112T imine reductase, conversion was observed toward (*R*)-salsolidine with 99% ee. Upon varying the AM, further cyclic imines or ketones could be reduced.



**Figure 15:** Representation of the double deracemization with MAO/artificial imine reductase (ATHase)/catalase, reproduced from Ward and co-workers.<sup>70</sup>

Interestingly, artificial imine reductases have also been applied in NADH regeneration within cascade reactions. NADH is an expensive cofactor and thus its regeneration has been investigated intensely.<sup>71,72</sup> Ward and co-workers used [Cp\*Ir(biot)Cl] – WT Sav coupled with the 2-hydroxybiphenylmonooxygenase (hbpA) and formate as the hydride source to regenerate NADH cofactor.<sup>70</sup>

Furthermore, Ward group teamed with Hollmann and co-workers to regenerate NAD mimics in a cascade using formate as hydrogen source.<sup>73</sup> The regeneration cascade was tested on five NAD mimics, using the iridium biotinylated-Sav coupled with an ene-reductase of the Old Yellow Enzyme family (*TsOYE*). Only one mimic reduced with enantioselectivity (up to 98 % ee-(*R*)) and high turnover number (100-1000 TON).

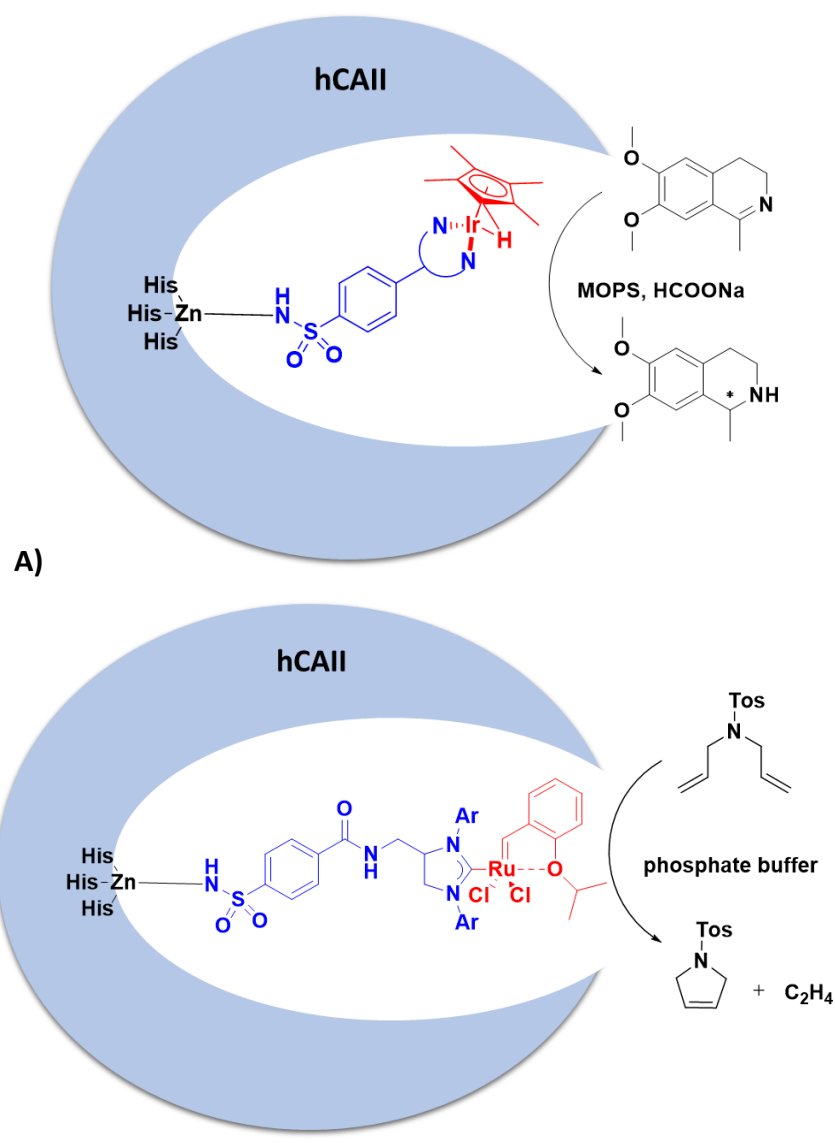
Subsequently, Ward and co-workers developed a cascade of reactions *in vivo*, with the biotin-(strept)avidin technology. Using their library of artificial imine reductases, they screened for an AM able to use the natural NAD(P)H cofactor as the hydride source for transfer hydrogenation reactions (a critical parameter to work *in vivo*).<sup>74</sup> The authors engineered several *in vivo* cascades with natural enzymes and catalysed several reaction types (e.g. hydride transfer reaction, olefin metathesis).<sup>75</sup> As an example, they designed a gene switch inside a mammalian cell that was activated by a molecule produced with their AM (engineered HEK-293T mammalian cells).<sup>76</sup> The gene activation then upregulated the expression of a bioluminescent nanoluc marker. One challenge of the system was the required integration of AM inside the

cells. The authors optimised their AM with a cell penetrating poly(disulfide) (CPD), in order to maximise the uptake of the AM. After addition of AM substrate, the luminescence in the presence of both AM-(CPD) and Hek-293T cells was significantly superior to the cells alone or to the cells with AM without CPD. Ward and co-workers highlighted the importance of cell penetration for AM used *in vivo*. They later succeeded in showing the expression and the activity of AMs in the periplasm<sup>77</sup> and at the surface of *E. coli*.<sup>78</sup>

Following the success of biotin-(strept)avidin technology and to increase the scope of scaffolds, other proteins have been studied for supramolecular anchoring. In a similar concept, the identification of high affinity between the scaffold and the potential ligand anchor is essential. For example, Mahy and Ricoux group studied the neocarzinostatin scaffold, naturally evolved to strongly bind testosterone. They coupled the testosterone with Fe(III)<sup>79</sup> Zn(II)<sup>80</sup> and Cu(II)<sup>81</sup> to create AMs catalysing respectively the oxidation of thioanisole, the hydrolysis of phosphate ester and the Diels-Alder cyclisation reactions. However, in all these models, low enantioselectivity was observed.

Ward and co-workers also worked successfully on another type of host protein: human carbonic anhydrase II (hCAII). The protein is known to bind aryl sulfonamides cofactors. The authors bound several sulfonamide-substituted Ir d<sup>6</sup> piano-stool catalysts [(η<sup>5</sup>-Cp\*)Ir(pico)Cl], to generate artificial transfer hydrogenases, with weak enantioselectivity (68 % ee-(S) salsolidine, **Figure 16A**).<sup>82,83</sup> Later, through computational modelling using Rosetta Design, they genetically optimised their AM model based on the crystal structure (PDB 3ZP9).<sup>84</sup> This resulted in a more tightly bound catalyst by new packing interactions and a stabilisation of the protein backbone through H-bonds. The catalyst was embedded in a fixed orientation within the protein. Experimental tests showed an increase in the affinity of [(η<sup>5</sup>-Cp\*)Ir(pico)Cl] for the enzyme in all the new designs. Besides, activity and enantioselectivity were significantly improved from the first model (up to 92 % ee-(S) salsolidine). In a similar rational design method, Ward and co-workers also introduced a Grubbs-Hoveyda catalyst to engineer artificial metathesases (**Figure 16B**).<sup>85</sup>





**Figure 16:** Two AMs adapted from Ward and co-workers. **A)** AM for the transfer hydrogenation of salsolidine. The bidentate ligand (blue) anchor the IrCp\* metal complex inside hCAII.<sup>83</sup> **B)** AM for ring-closing metathesis. The arylsulfonamide ligand (blue) is anchoring the Grubbs-Hoveyda type catalyst inside hCAII.<sup>85</sup>

Following a similar strategy to biotin-Sav, Tanaka and co-workers developed an AM based on the supramolecular interaction albumin-coumarin. They linked a Ru catalyst to a coumarin anchor, binding strongly inside the albumin hydrophobic pocket. The anti-cancer therapeutic potential was highlighted along with the interesting biocompatibility of this human based AM.<sup>86</sup>

The supramolecular anchoring strategy is a dynamic and self-adjustable system, conducive to wide optimisation opportunities. The different designs of AMs proved how changing the active site around the metal (second coordination sphere) impacted the activity (e.g. charged residues might decrease imine reductases

activity). Modifications of the features in direct coordination with the metal (first coordination sphere) also impact activity and selectivity (e.g. arene capping). A tightly bound catalyst seemed to react better than a loose one. All these conclusions showed how the design of non-covalent AMs can be challenging. Any subtle changes that can occur in the contacts between the anchoring ligand and the scaffold must be taken into account. These previous works also bring forward how understanding the structure of the scaffold and the position of the metal catalyst is critical.

### 1.3 Alcohol dehydrogenases

A critical requirement in the design of AMs is the selection of a suitable enzyme scaffold to interact with the metal catalyst and the substrate during the catalytic reaction. This project will focus on the reduction of carbon-heteroatom double bonds by modification of natural reductases.

#### 1.3.1 *Background*

One of the most prevalent biocatalysts for the reduction of carbonyls is the nicotinamide-dependent alcohol dehydrogenase (ADH),<sup>87</sup> present in many organisms (human, animals, bacteria, yeast...). ADHs detoxify alcohol in the mammalian organisms, and they are part of the fermentation process in yeast and bacteria. Another interesting fact about ADHs is the broad range of substrates they accept.

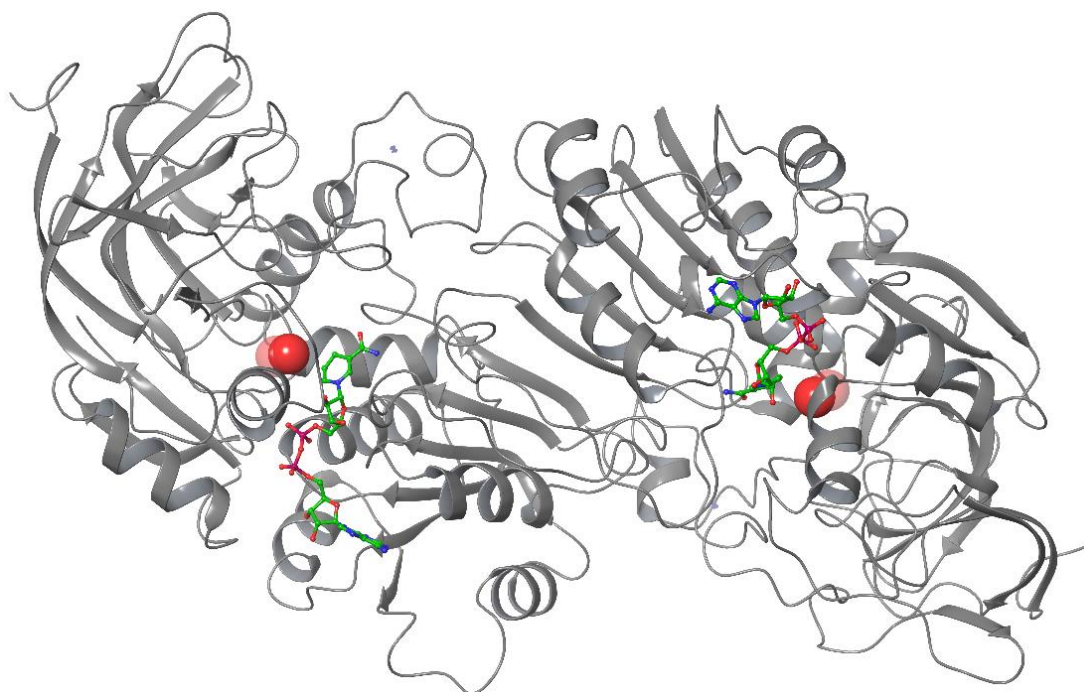
ADHs have been classified in three subfamilies depending on their amino acid chain length and structure: The short-chain dehydrogenases / reductases (SDRs), the medium-chain dehydrogenases / reductases (MDRs) and the long-chain dehydrogenases / reductases (LDRs). The common structural feature to all ADHs is the Rossmann-fold element, defined by a three-layered arrangement of a beta sheet composed of six beta strands surrounded by two alpha helices. This strand topology is the typical binding site for the NAD(P)H natural cofactor.<sup>88</sup>

ADHs catalyse the reversible oxidation of primary and secondary alcohols to aldehydes and ketones *via* transfer hydrogenation reaction, using the cofactor NAD(P)H / NAD(P)<sup>+</sup> as an electron acceptor/donor. The presence of natural, stable, and operative cofactor-enzyme interactions is interesting for the supramolecular anchoring strategy.

The two ADHs used as scaffold in the project are part of the MDRs sub-family: the horse liver alcohol dehydrogenase (HLADH) and the *Thermoanaerobacter brockii* ADH (TbADH). Both are zinc-dependent enzymes.

### 1.3.2 Horse liver alcohol dehydrogenase (HLADH)

The HLADH crystal structure has been resolved since 1976 by Eklund and co-workers (PDB 4XD2, **Figure 17**).<sup>89</sup> Mammalian ADHs are dimer structures, built on two domains: the C-terminal domain, composed of the Rossmann fold motif (cofactor binding site), and the N-terminal catalytic domain for the substrate binding. Two zinc ions were identified in the HLADH structure. The first  $\text{Zn}^{2+}$  is coordinated by four cysteines (C97, C100, C103 and C111) and is used for the structural stabilisation of the enzyme. The second zinc ion is a catalytic  $\text{Zn}^{2+}$ , tetrahedrally-coordinated by C174, C46, H67 and a molecule of water which is replaced by the hydroxyl group of the alcohol substrate during catalysis. This  $\text{Zn}^{2+}$  ion is located at the intersection of the cofactor and the substrate binding pocket.<sup>90,91</sup>

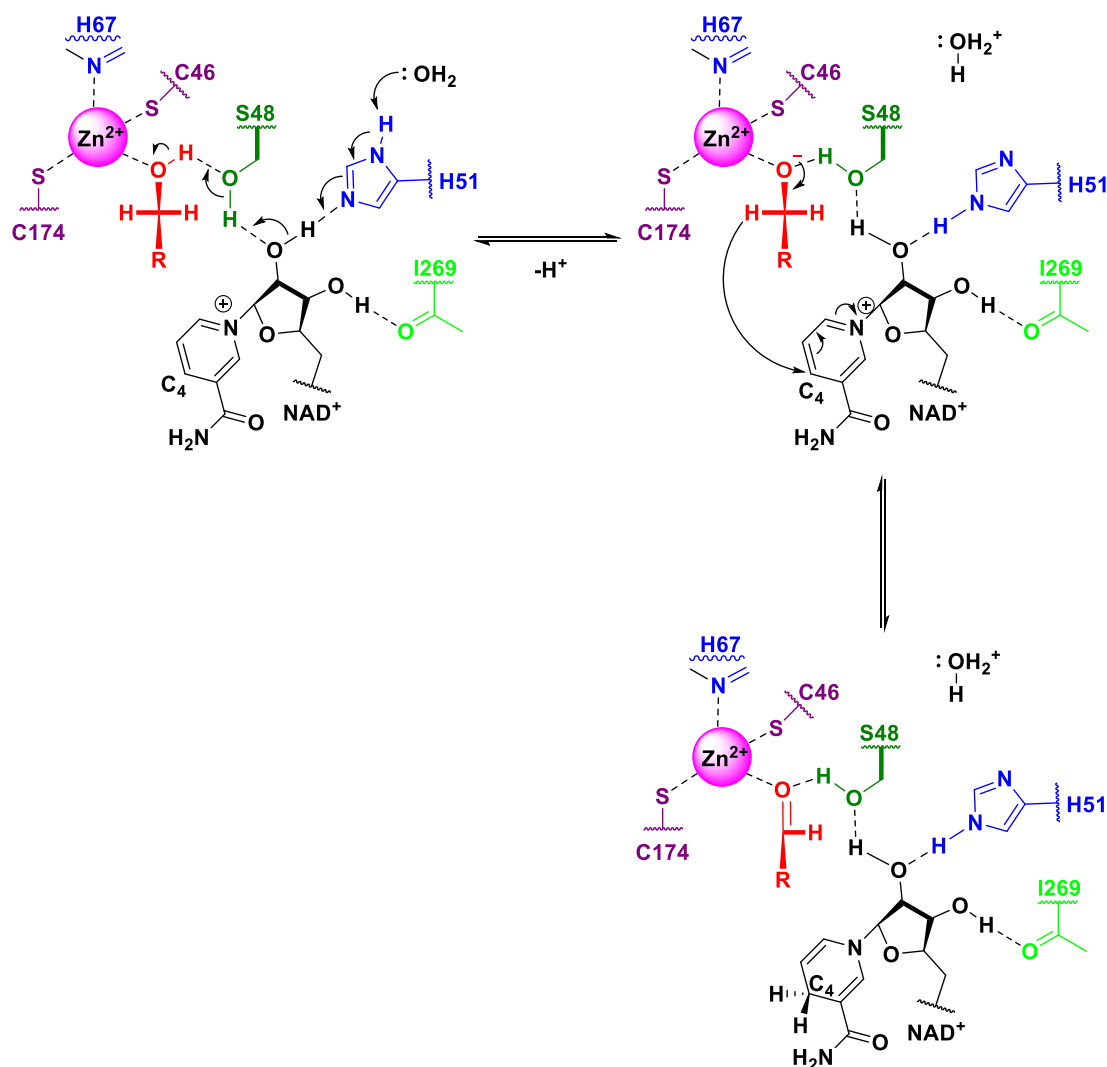


**Figure 17:** Horse liver ADH (HLADH) structure (PDB 4xd2) with NAD(H) cofactor (green) at the domain's interfaces.  $\text{Zn}^{2+}$  are displayed in red spheres.<sup>92</sup>

#### 1.3.2.1 Mechanism of action

The catalytic mechanism of HLADH implies a hydrogen transfer between the natural  $\text{NAD}^+$  cofactor and the alcohol substrate (**Figure 18**). The oxidation reaction starts by

the binding of NAD<sup>+</sup> into the cofactor pocket,<sup>91</sup> followed by the alcohol substrate binding to Zn<sup>2+</sup> *via* displacement of a molecule of water. An extended connection is then established: the 2-OH of the nicotinamide ribose is hydrogen bonded to H51 which engages a cascade of proton transfer to deprotonate the zinc-bound alcoholate and transfer the hydride toward the nicotinamide ring of NAD<sup>+</sup>.



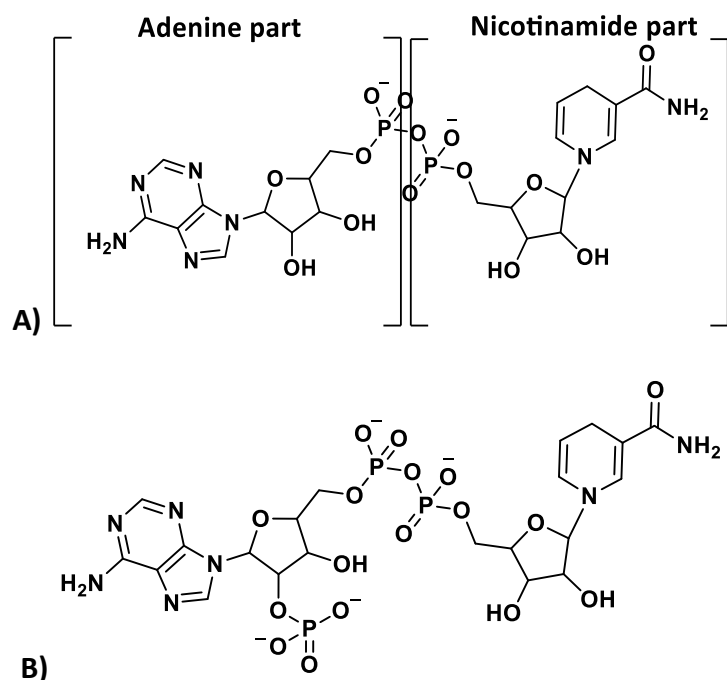
**Figure 18:** Catalytic mechanism of HLADH based on proton-relay, reproduced from Lee and co-workers.<sup>93</sup>

After the dissociation of the aldehyde product from Zn<sup>2+</sup> and replacement by a water molecule, the reduced NADH form leaves with the opening of the cofactor binding pocket.

The reverse hydride transfer reaction is also observed from the reduced NADH cofactor to the carbonyl carbon of an aldehyde, liberating the NAD<sup>+</sup> oxidized form.<sup>94</sup>

### 1.3.2.2 Cofactor binding

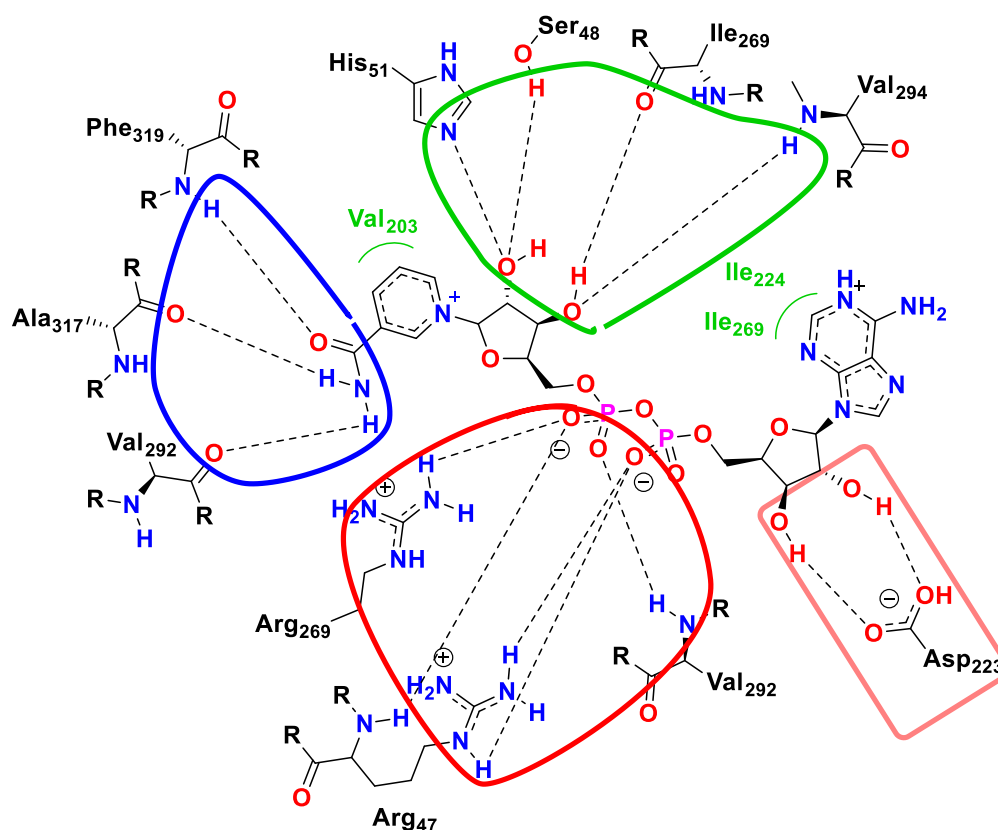
The different ADHs possess distinct specificities for their natural cofactor. Some are more active with NAD(H), one of the most important cofactors found in cellular mechanisms (e.g. gene expression, cell death, signal transduction),<sup>95</sup> while others prefer NADP(H) (**Figure 19**). The human liver ADH and HLADH use NAD(H).<sup>96</sup>



**Figure 19:** Structure of the ADHs natural cofactors **A)** NADH and **B)** NADPH. The structure is divided in two parts: the adenine ring and the nicotinamide ring, used for the catalytic activity. The two parts are linked by a diphosphate group. Their oxidized forms are respectively  $\text{NAD}^+$  and  $\text{NADP}^+$ .

Upon the binding of NAD(H) cofactor, a drastic change in domain conformation closes the cofactor binding site and brings NAD(H) closer to the catalytic site.<sup>97</sup>

Inside the narrow cofactor binding pocket, the adenine of NAD(H) is placed in a hydrophobic cleft, pointing toward the solvent. The ribose hydroxyls are hydrogen bonded to an acid aspartic residue side chain, the D223 (**Figure 20**). The middle pyrophosphate group is bound through its oxygens to two arginines side chains. The nicotinamide is composed of one pyridine ring (N-ring), important for the catalytic reactivity, and an amide group interacting with F319, A317 and V292 residues. Hydrophobic interactions between one face of the N-ring and V203, V292 and T178 residues place the nicotinamide part close to the active site and  $\text{Zn}^{2+}$ .<sup>91</sup> The correct binding position of the nicotinamide ring is particularly important for the catalytic activity.



**Figure 20:** Schematic representation of binding interactions between HLADH crystal structure (PDB 4xd2) and NAD<sup>+</sup>: amide binding (blue), nicotinamide ribose binding (green), pyrophosphate binding (red) and adenine ribose binding (orange).

The specificity of ADHs for the two cofactors has been explained in several studies by comparison of catalytic activity. However, the catalytic activity only gives an indication of binding but not a precise affinity strength or insights in the binding process.

Fan and co-workers showed that the presence of D223, in the active site of HLADH acts as a charge repulsive for NADPH.<sup>98</sup> The presence of a phosphate on the ribose just in front of the aspartic acid residue leads to an electrostatic repulsion, preventing the correct position of NADPH. To support their theory, the corresponding amino acid (D201) in the yeast ADH I was mutated into a glycine. They realised steady-state kinetic studies on ethanol oxidation. The results highlighted changes toward an equal use of NAD<sup>+</sup> and NADP<sup>+</sup>. The authors suggested that D223 was responsible of the exclusion of NADP(H) cofactors.

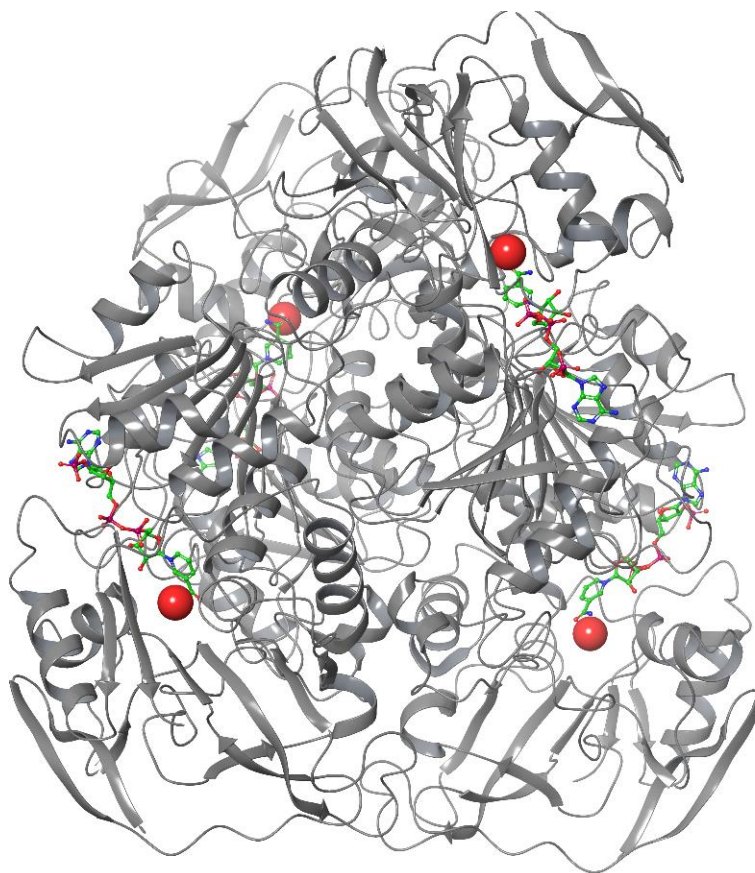
More recently, a study compared HLADH to the complex NADPH-ADH8 (ADH from amphibian *Rana perezi*), the only vertebrate ADH specific to NADPH.<sup>99</sup> In ADH8, contacts along the binding pocket remain identical to HLADH at the exception of a

cleft in place of D223. The small pocket is composed of neutral and basic amino acids (G222, T223, H224, K227 and L199), allowing the insertion of the NADPH phosphate. This study confirms again Fan's conclusions that D223 is responsible for the NADH specificity in ADHs. The presence of the phosphate group in NADP(H) leads to electronic and space clashes inside liver ADH, probably reducing the binding affinity and thus the activity.

### 1.3.3 *Thermoanaerobacter brockii* alcohol dehydrogenase (TbADH)

Tetrameric type of ADHs can also be found, mainly in fungal, bacterial or yeast enzymes.<sup>96</sup> TbADH is a tetrameric alcohol dehydrogenase from the bacteria *Thermoanaerobacter brockii*, first described in 1980 by Lamed and Zeikus (**Figure 21**).<sup>100</sup> Each monomer are composed of a Rossmann fold cofactor binding site connected to a catalytic site by an  $\alpha$ -helix.<sup>97</sup>

Like most of the alcohol dehydrogenase, TbADH activity depends on a  $\text{Zn}^{2+}$  ion located in the catalytic site. In contrast to HLADH,  $\text{Zn}^{2+}$  is bound to four amino acids: C37, H59, E60 and D150 (replaced by a molecule of water in HLADH).<sup>93</sup>

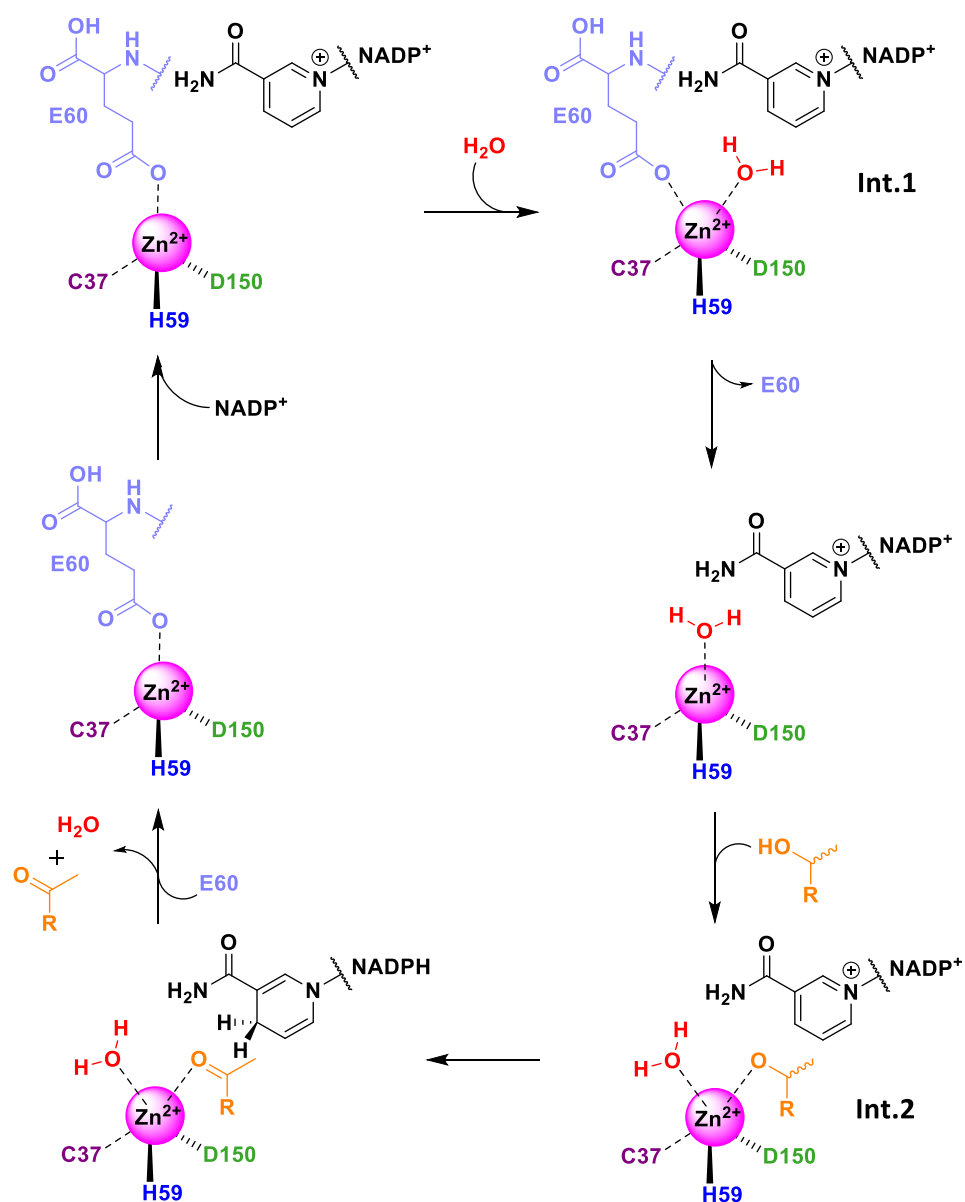


**Figure 21:** Holo-TbADH crystal structure at 2.5 Å (PDB 1YKF) with its natural cofactor (green) and the  $\text{Zn}^{2+}$  ions in red spheres.<sup>97</sup>

TbADH is also well known for its high stability: the enzyme is stable at high temperature (up to 93 °C) and shows resistance to organic solvents.<sup>97</sup>

### 1.3.3.1 Mechanism of action

The TbADH mechanism follows an alternative model to the one explained in HLADH (**Figure 22**). The catalytic  $\text{Zn}^{2+}$  forms two penta-coordinated complexes with water and substrate, and no extended proton relay transfer are necessary.<sup>93</sup>



**Figure 22:** Reproduction of the catalytic mechanism of TbADH from Lee and co-workers. **Int.1** and **Int.2** are the two intermediates penta-coordinated.<sup>93</sup>

The binding of NADP<sup>+</sup> forms the first penta-coordinated zinc complex, with the addition of a water molecule (**Int.1**). A tetra coordinated  $\text{Zn}^{2+}$  ion is back with the

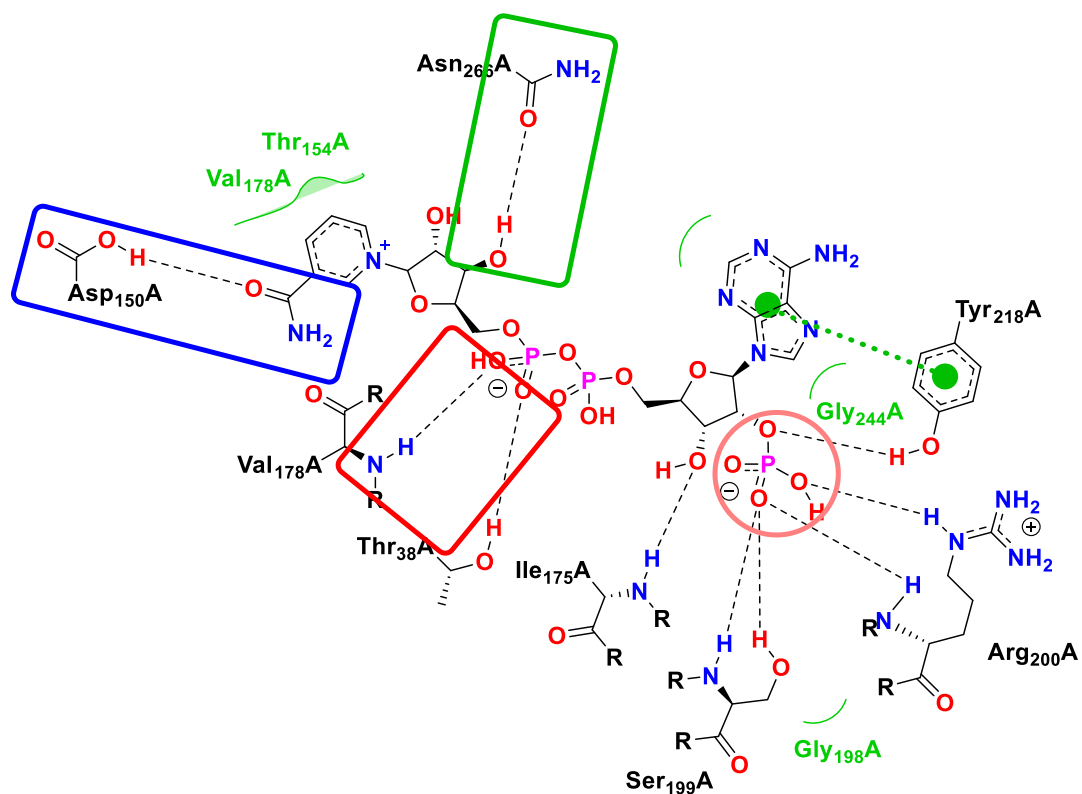


dissociation of E60. The subsequent binding of the alcohol substrate leads to the second intermediate penta-coordinated (**Int.2**). The hydride transfer is then completed and yield the cofactor reduced form NADPH and the aldehyde/ketone product. The detachment of the product induces the association of E60 to provide the original tetra-coordinate zinc complex.

#### 1.3.3.2 Cofactor binding:

Only 27 % of TbADH sequence is identical to eukaryotes enzymes, including HLADH. The majority of the similarities is concentrated in the cofactor binding site (43 % identity).<sup>101</sup> Nonetheless, TbADH is among the few ADHs dependent to NADP(H). The specificity for NADP(H) cofactors over NAD(H) is explained in the binding site composition and support the previous explanations of HLADH specificity (section1.3.2.2, page 24).

Frolow and co-workers affirmed the specificity for NADP(H) cofactor relies on four residues: G198, S199, R200 and Y218 (**Figure 23**).<sup>97</sup> First, the G198 replaces the D223 residue found in all NAD(H)-dependent enzymes, including HLADH. As discussed previously, this aspartic acid creates charge repulsion with the phosphate ribose of NADP(H). G198 residue is composed of a smaller side chain, fitting the extra phosphate. The three other residues are then making hydrogen bonds with the extra phosphate of NADP(H). Through the adenine part binding, the whole cofactor is stabilised and more efficient.



**Figure 23:** Interactions of TbADH (PDB 1ykf) with NADPH: amide binding (blue), nicotinamide ribose binding (green), pyrophosphate binding (red) and adenine phosphate ribose binding (orange).<sup>97</sup>

Upon binding of the cofactor, the four catalytic domains of the monomers initiate a movement that slightly changes conformation.<sup>97</sup> As a consequence, the whole tetramer structure expands. In contrast, HLADH is subject to more drastic changes of the rigid body, narrowing the pocket.

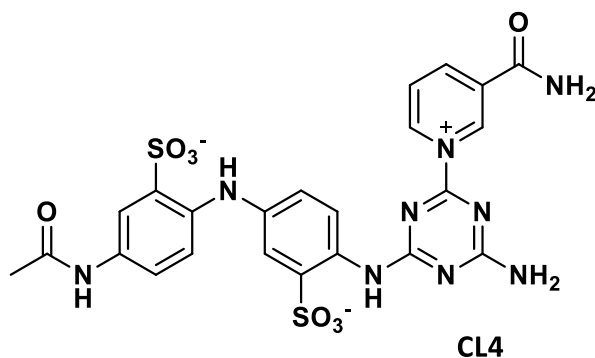
The cofactor binding diagram clearly shows less hydrogen bonds than in HLADH one (**Figure 20**). Here, only one amino acid binds to the amide and the ribose hydroxyl, and only two residues exchange hydrogen bonds with the pyrophosphate. The more intensive binding is held by the adenine part, explaining again the specific affinity of TbADH for NADP(H) cofactor.

#### 1.3.4 Cofactor analogues

The high cost of NAD(P)H cofactors have led studies to seek for alternative, less expensive NADH mimics.<sup>102</sup> Generally, the mimics were designed with the replacement of the adenosine-phosphate-ribose part.

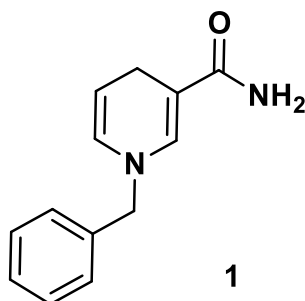
At first, Ansell and co-workers designed a cofactor analogue called CL4, based on the nicotinamide ring substituted by triazines (**Figure 24**).<sup>103</sup> They tested CL4 inside

HLADH for the oxidation of butan-1-ol. Compared to  $\text{NAD}^+$  cofactor, the authors obtained a lower efficiency (low  $V_{\text{max}}$  and high  $K_M$ ).



**Figure 24:** CL4 mimic is an analogue of NADH cofactor, designed by Ansell and co-workers. Only the nicotinamide ring was preserved and substituted by triazines.<sup>103</sup>

Later, Fish and co-workers synthesised an even more truncated NADH biomimetic, with a benzyl group in place of the adenine, pyrophosphate and ribose parts (**1**; **Figure 25**).<sup>104,105</sup> Several mimics were then synthesised, which kept the same model: a N-ring substituted by a benzyl group in place of adenine and pyrophosphate parts.



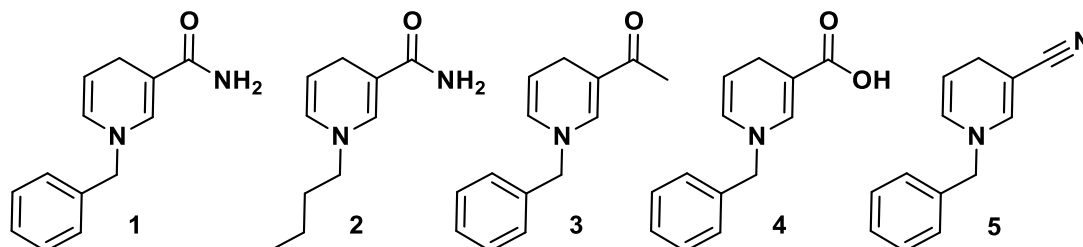
**Figure 25:** 1-Benzylnicotinamide, used by Fish and co-workers to mimic NADH cofactor.

The authors implied a recognition by HLADH and a catalytic activity of these mimics, using the reduction of several achiral ketones as model. For example, a yield of 90 % with 93 % ee was obtained with the mimic **1**, while  $\text{NAD}^+$  obtained 91 % yield and 93 % ee. They thus concluded that only 1,4-dihydronicotinamide part is necessary in the recognition and the binding to the cofactor binding site.

However, these results have since been challenged by Paul and Hollmann, who linked the reactivity of these  $\text{NAD}^+$  mimics to the presence of NADH inside a non-pure HLADH protein.<sup>106</sup> They later used a purified HLADH with the similar synthetic mimic **1** and assumed from their preliminary results that the enzyme cannot bind the mimic. The authors suggested that some functional groups of NADH which were truncated

(mainly ribose or pyrophosphate) might play a relevant role in the affinity toward HLADH, and therefore the catalytic activity.

Hollman group later investigated several NADH mimics in a flavin-dependent ene-reductases (ER) for ketone and aldehyde reduction (**Figure 26**).<sup>107</sup> Their results were equivalent to those with the natural NADPH cofactor, or even better.



**Figure 26:** NADH mimics reduced form, tested inside ER enzymes by Hollman and co-workers for oxidoreduction reactions.<sup>107</sup>

However, attempts with the same mimics on HLADH ended with no activity detected.<sup>108</sup>

To the best of our knowledge, no literature evidence shows NAD(P)H cofactor mimics assays inside TbADH. The only interactions of mimics inside TbADH are found in Dr Mattias Basle's thesis.<sup>109</sup> He explored the binding of cofactor mimic **1** and its oxidised form inside both HLADH and TbADH, through activity and inhibition tests. The conclusions displayed no interaction between both enzymes and the derivatives.

### 1.3.5 Conclusion

Artificial metalloenzymes is a prosperous strategy in biocatalysis. The engineering of these catalysts depends on multiple factors, among which the selection of a proper enzyme scaffold to stabilise the metal catalyst and the substrate in a catalytic orientation. In this project, the transfer hydrogenation of imines was selected as the chemical reaction targeted for the AM.

Zn<sup>2+</sup>-dependent ADH enzymes activate the reduction of carbonyls by hydrogen transfer reaction, but despite their structure similarity with ketones, imines do not get reduced by ADHs. The stability, the structure, and the large substrate scope of ADHs make them interesting for the design of AMs.

In addition, the natural NAD(P)H cofactor has already evolved to tightly bind inside ADHs, which represents an advantage for the supramolecular binding of a metal catalyst. For now, the literature indicates a lack of activity of short synthetic mimics

inside ADHs. This poor activity may be the result of a lack of crucial bounds inside the cofactor binding site, however this has never been confirmed by affinity studies in ADHs. A strong knowledge of NAD(P)H binding mode is thus required to improve the design of synthetic cofactors and increase affinity for ADHs.

## 1.4 Natural and synthetic hydrogen transfer reactions

The second requirement for the engineering of an AM is the choice of the metal catalyst, in agreement with the project objective: the modification of the previously described natural ADH for the transfer hydrogenation of carbon-heteroatom double bonds.

### 1.4.1 *Background*

The reduction by hydrogen transfer consists of an intermolecular transfer of hydrogen with its electron pair to an activated carbon. This process results in an oxidoreduction. These reactions are frequent in nature, where enzymes use organic hydride sources in cofactors such as the NADH or flavin adenine dinucleotide (FAD).<sup>110</sup> Nature inspired chemists to create new organocatalysts, metal catalysts and biocatalysts for industrial applications. In 1903, Bergdolt and co-workers first reported the hydrogen transfer between two molecules catalysed by palladium.<sup>111</sup> The challenges of this reaction are the design of an efficient catalyst with enhanced yield and better enantioselectivity.

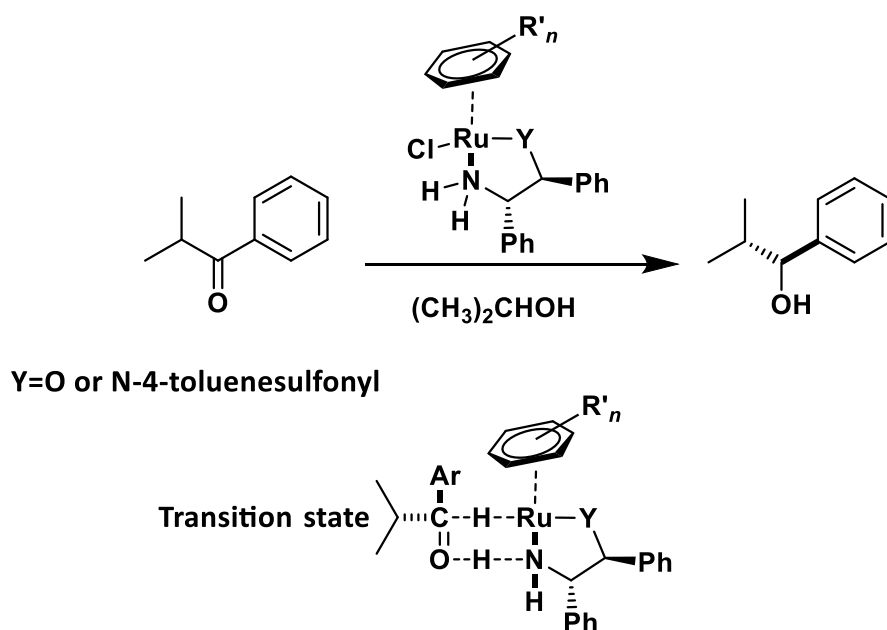
### 1.4.2 *Chemical catalysis: Organocatalysts and metal catalysts*

#### 1.4.2.1 Organometallic catalysts

In 1925, the first metal-catalysed transfer hydrogenation of a carbonyl compound was developed independently by Meerwein and Schmidt<sup>112</sup> and by Verley.<sup>113</sup> They reduced ketone and aldehyde in their corresponding alcohol using an aluminium alkoxide (now called the Meerwein–Ponndorf–Verley reduction reaction).

Later, increased activity was discovered using transition metal catalysts. Rh, Ru and Ir have been extensively studied for their reactivity. Notably, Mason and co-workers used Ir complexes to catalyse the reduction of cyclohexanone to the corresponding alcohol, using water as the hydrogen source.<sup>114</sup> Besides, Sasson and Blum developed Ru complexed with phosphine to successfully catalyse the transfer hydrogenation of ketones.<sup>115,116</sup> However, in all these models, enantioselectivity was still a challenge and the production of chiral compounds required difficult purification steps.

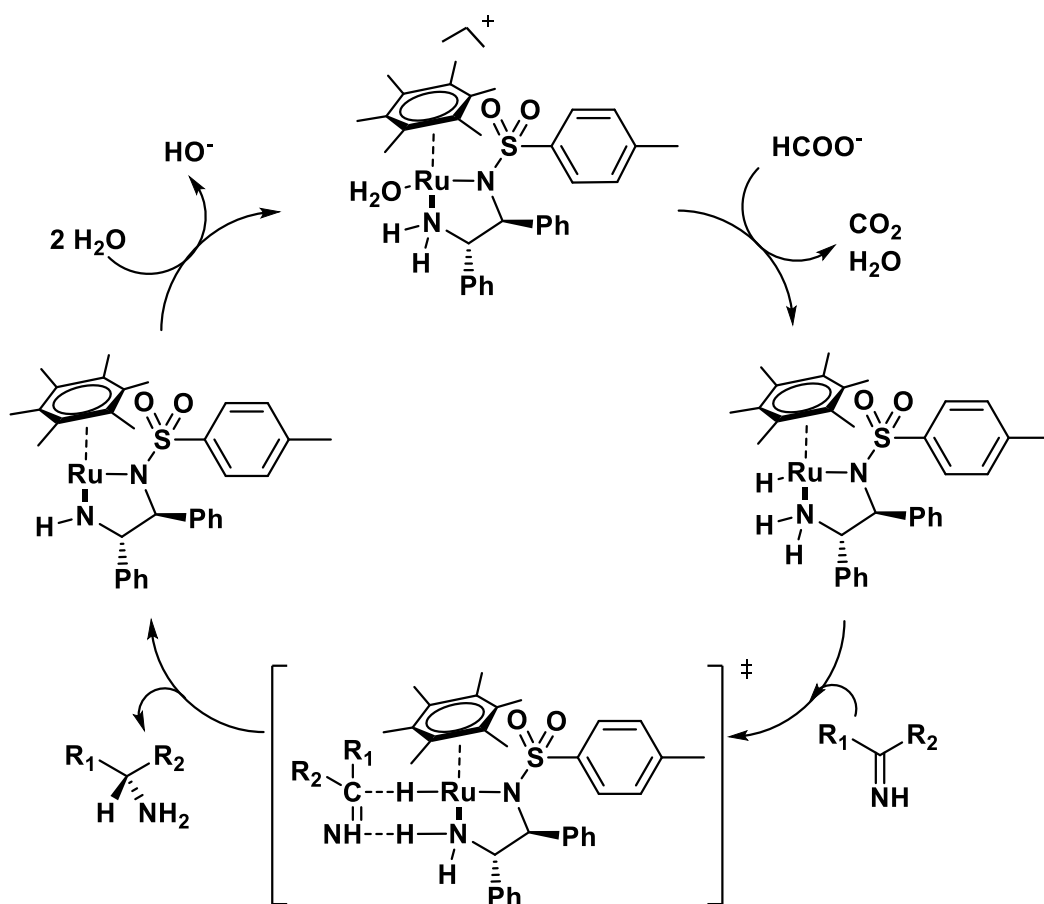
The major breakthrough in asymmetric hydrogenation came from Knowles<sup>117</sup> and Noyori,<sup>118</sup> separately developing chiral complexes showing enantioselectivity on carbonyl and imines reduction by hydrogen transfer (respectively Rh(DIPAMP) and Ru-BINAP). Noyori and co-workers proved the transfer hydrogenation of ketone using Ru-d<sup>6</sup> piano-stool complexes, but also the versatility of the catalyst by catalysing the oxidation of secondary alcohol (**Figure 27**).<sup>56</sup> Later, Noyori reported the first imine reduction in 1996, by developing a piano-stool  $\pi$ -complex [Ru-TsDPEN] (TsDPEN= N-(*p*-toluenesulfonyl)-1,2-diphenylethylenediamine) in combination with formic acid / Hünig's base as hydrogen source.<sup>119</sup> The importance of the  $\eta^6$ -arene and the diamine ligand for the high selectivity and the broad substrate scope was proven.<sup>120</sup>



**Figure 27:** Asymmetric transfer hydrogenation of carbonyl, catalysed by Noyori's chiral Ru complexes. The metal complex favoured transition state is depicted underneath the reaction scheme. Adapted from Noyori's work.<sup>118</sup>

The development of these transition metal catalysts started the race toward more efficient chiral selectivity of products for industrial applications. The most successful catalysts are based on Ru(II), Rh(III) and Ir(II) half-sandwich  $\pi$ -complexes, with a chiral bidentate ligand and an arene or Cp\* ( $\eta^5$ -pentamethylcyclopentadienyl) cap.<sup>121</sup> The success of these catalysts comes from the achiral ligand for selectivity and the transition metal ion in low oxidation state for the control of activity. The hydrogen source comes from molecular hydrogen or organic hydrogen (e.g. formic acid, 2-propanol).

Recent studies focus on a more environmentally friendly use of these transition metal complexes. Chiral arene Ru, Rh and Ir catalysts have been reported to catalyse the transfer hydrogenation of ketone and imine in aqueous media.<sup>122,123,124,125</sup> A tosyl substituted chiral diamino ligand coordinates the transition metal ions (**Figure 28**). Sodium formate is used as the source of hydrogen. When high quantities of formate are used, the catalytic rate is increased.



**Figure 28:** Hypothetic catalytic mechanism for the enantioselective transfer hydrogenation of imine, using a biotinylated Ru bidentate catalyst in aqueous solution with sodium formate as hydride source. Adapted from Suss-Fink and co-workers postulated mechanism.<sup>124</sup>

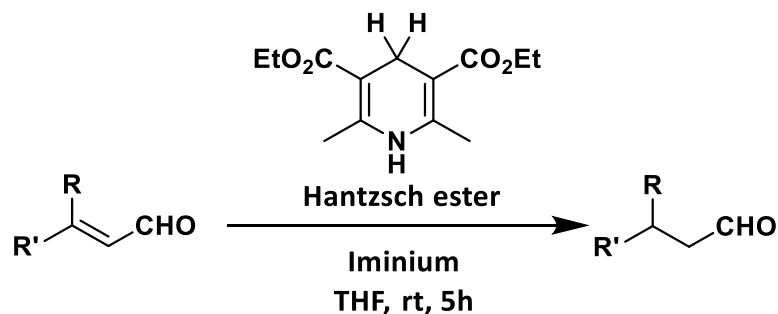
Even closer to nature, examples from Ward<sup>74</sup> and Soetens<sup>126</sup> use similar metal catalyst with NADH cofactor as the hydrogen source. More recently, the change of metal to Osmium resulted in the reduction of pyruvate *in vivo*, in cancer cells.<sup>127</sup>

#### 1.4.2.2 Organocatalysts

For greener perspectives, organocatalysts are an ecological alternative to the use of metal ions in transfer hydrogenation reactions.<sup>110</sup>



List and co-workers<sup>128</sup> with MacMillan and co-workers,<sup>129</sup> reported separately the first metal-free transfer hydrogenation reaction (**Figure 29**). They used several ammonium salt as efficient catalysts, in combination with the Hantzsch ester as a hydrogen source mimicking the NADH cofactor.



**Figure 29:** Transfer hydrogenation of double bonded carbon, catalysed by iminium combined with Hantzsch ester, separately reported by List and MacMillan groups.<sup>128</sup>

Both groups later introduced asymmetry in a variety of their reactions.<sup>130,131,132</sup> By introduction of chiral phosphate anion, List group proved a high enantioselectivity with a large scope of substrates containing C=C and C=N. They later extended their reaction scope toward the reduction of carbonyls by re-introducing the metal catalyst Cu(II)-bisoxazoline.<sup>133</sup>

Even if less reactive than transition metal catalysts, organocatalysts are sometimes more stable than enzymes or organometallics, like in solid state reactions.<sup>110</sup>

#### 1.4.3 Biocatalysis: Transition metal catalysts within a protein scaffold host

##### 1.4.3.1 C=C double bond reactions

Enoate-reductases (ER) are the preferred natural catalysts for the C=C reduction by hydrogen transfer.<sup>87</sup> However, their dependence on expensive NAD(P)H cofactors represents a major challenge for industrial applications, which is balanced by additional efforts to externally recycle the cofactor. New alternatives are the development of AMs.

Hollmann and co-workers<sup>107</sup> and Qian and co-workers<sup>134</sup> created mimics of NADH inside the Old Yellow Enzyme family, to replace the expensive NAD(P)H in the catalysis of hydrogenation reactions (examples in **Figure 26**, page 31).

Ward and co-workers reported on the hydrogenation of acetamidoacrylic acid using Sav AMs, from a library of Rh(diphosphine)(cod) biotinylated catalysts.<sup>54,53</sup> The study

showed the importance of the metal catalyst structure. A more flexible linker between the metal and biotin catalysed the hydrogenation with better enantioselectivity than more rigid linkers.

#### 1.4.3.2 Carbonyl (C=O) reduction reactions

Dehydrogenases are the preferred enzymes to reduce carbonyls. Their large substrate scope and their high enantioselectivity (frequently surpassing chemical catalysts) make them the prime choice of chemical substitute.<sup>87</sup> However, as seen in ERs, their dependence on the expensive cofactor NAD(P)H is a bias in their use at big scale. Alternatives rely again on recycling the cofactor or generating artificial dehydrogenases with low-cost catalysts.

Noyori's catalyst has been inserted within diverse AMs for the transfer hydrogenation of carbonyl. Salmain and co-workers used papain and bovine  $\beta$ -lactoglobuline proteins as host for several d<sup>6</sup>-Rh-bipyridines complexes.<sup>135,136</sup> The resulting AMs catalysed the hydrogenation of cyclic ketones and NAD<sup>+</sup> cofactor.

A big step forward came from Ward and co-workers, who developed a now well-known AM based on Sav, for the efficient reduction of ketones.<sup>57</sup> They synthesised biotinylated catalysts composed of the aminosulfonamide Ru  $\eta^6$ -arene complex attached to the biotin molecule and introduced into Sav (**Figure 27**, page 34). This new AM proved to be efficient in the asymmetric reduction of ketones with high enantioselectivity (94 % (*R*)).

#### 1.4.3.3 Imine (C=N) reduction reactions

Chiral amines are important constituent of bioactive molecules derived from natural or chemical sources, with applications in chemical industries such as the fragrance, pharmaceutical or agricultural industry.<sup>137</sup>

An efficient route to chiral amines is the reduction of imines. Natural imine reduction is performed by imine reductases (IREDs), discovered by Nagasawa and co-workers<sup>138</sup> in 2011, which marked the beginning of their exploitation. Notably, the work of both Hauer and co-workers<sup>139</sup> and recently Turner and co-workers<sup>140,141,142</sup> extended the investigations on IREDs structures, substrate scope and activity. Specific (*S*) and (*R*) IREDs were identified, and their substrate explored.

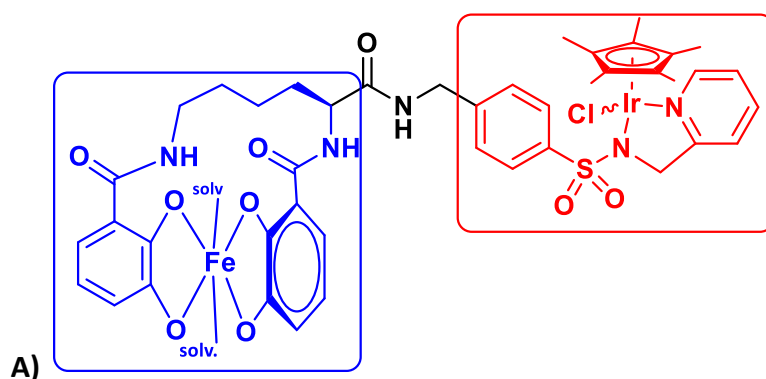
The best artificial imine reductase reported for the transfer hydrogenation of imine comes from Ward and co-workers work on biotin-streptavidin. Investigations on the chemical part (optimisation of ligands and metal centres) and on the genetic side (2<sup>nd</sup> coordination sphere mutations), improved the scope of the reaction.<sup>10</sup>

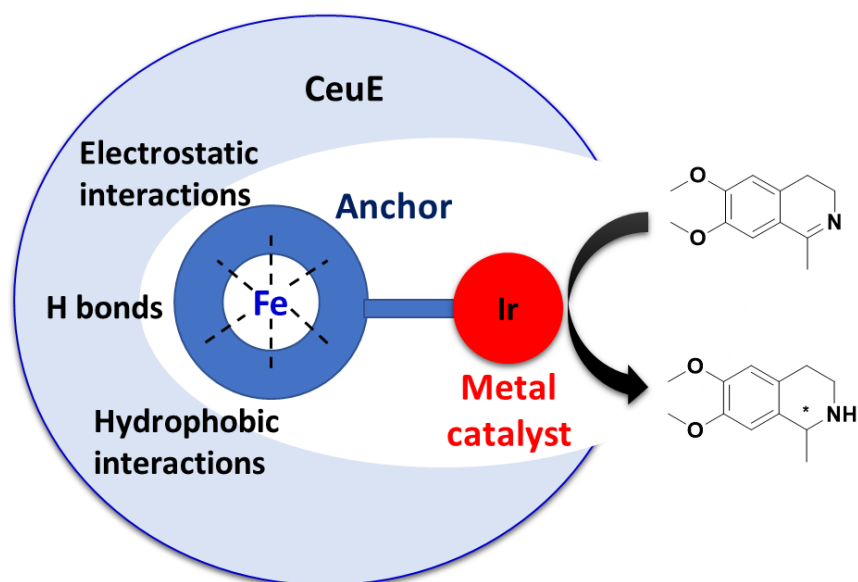
The authors identified the best complex for imine reduction as an Ir piano-stool inside streptavidin Sav112 mutant. Both (*R*)-salsolidine (95 % ee) and (*S*)-salsolidine (86 % ee) enantiomers were synthesised (**Figure 13**, page 16).<sup>64</sup>

Based upon this work, Rimoldi and co-workers validated the use of a different ligand (1,3 aminosulfonamide) for the development of imine reductases.<sup>143,144</sup> These AMs have then been applied to reaction cascade with natural enzymes, and coupled to the regeneration of NADH.<sup>70,74</sup>

In recent work, Ward and co-worker changed the scaffold for the asymmetric hydrogenation of ketones and imines by using the hCall. With the help of computational modelling, they built competent mutants of hCall with Ir complexes.<sup>84</sup>

Another strategy, by Duhme-Klair and co-workers, used a sulfonamide Ir(III)-complex catalyst inserted into the iron-siderophore periplasm binding protein (CeuE) of *Campylobacter jejuni*, to create an artificial transfer hydrogenase (**Figure 30**).<sup>145</sup> The iron-siderophore anchoring part allowed for a reversible attachment of the transition metal. Low efficiency was observed for the reduction of imine (36 % ee-(*R*)) but the AM assembly had the advantage to be reversible and controlled.





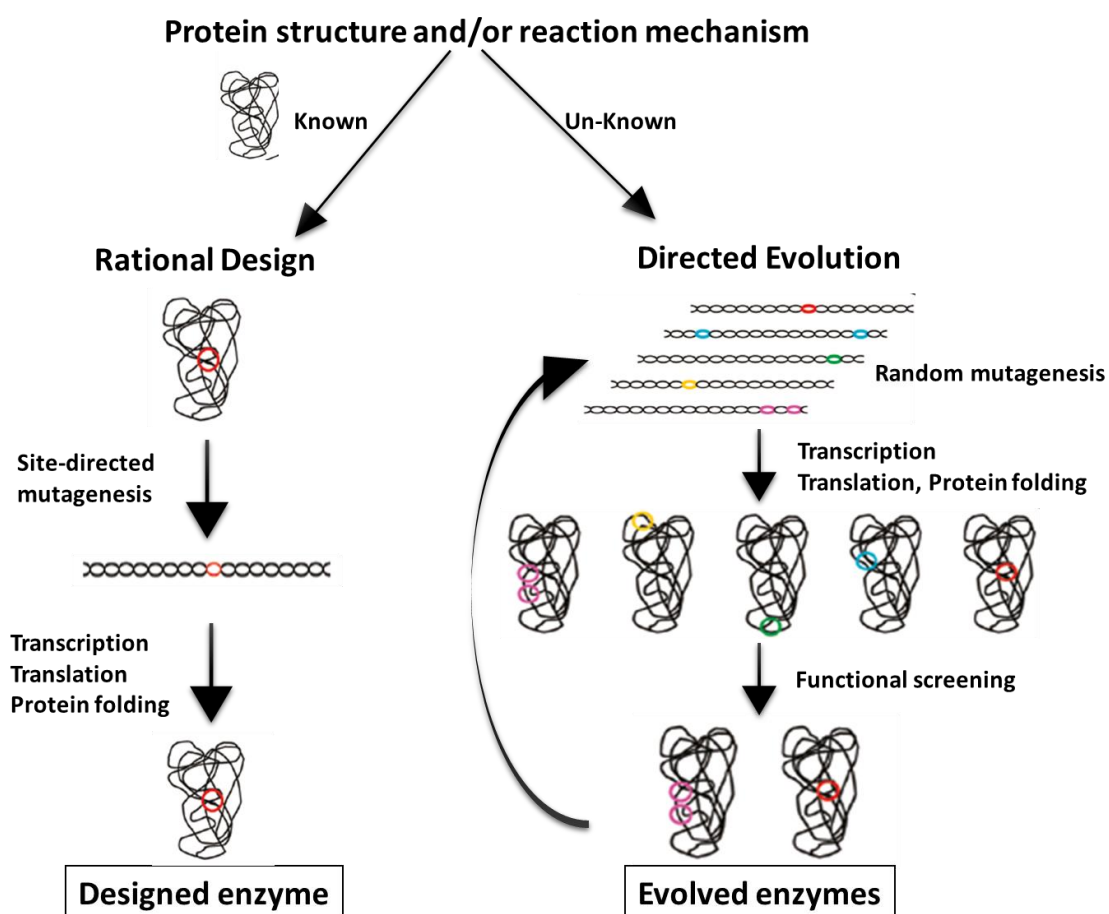
B)

**Figure 30:** Artificial transfer hydrogenase created by Duhme-Klair and co-workers, combining CeuE protein and a sulfonamide-Ir(II) complex. **A)** Catalyst engineered for the synthesis of salsolidine. In blue is highlighted the iron-siderophore ligand, and in red the Ir-complex catalyst. **B)** Design of the siderophore (blue) anchoring the iridium catalyst (red) inside CeuE. Adapted from Duhme-Klair and co-workers<sup>145</sup>

## 1.5 Enzyme engineering

### 1.5.1 Background

The biocatalysis field is in constant evolution with the improvement in protein engineering technologies, offering great possibilities for the design of tailored catalysts. The traditional design and optimisation of enzymes largely centred around two main approaches: directed evolution and rational design (**Figure 31**). The engineering of enzymes relied mostly on experimental studies. However, nowadays, the use of experimental procedures tends to be supplemented by computational techniques. Even though the methods still need improvements, computational tools can provide cheaper and quicker access to knowledge on structures and chemical reaction mechanisms.<sup>146</sup>



**Figure 31:** Scheme of two approaches of protein engineering: rational design and directed evolution. Adapted from Gruber-Khadjawi and co-workers representation.<sup>147</sup>

Computational protein design has been applied either to re-design native proteins towards different substrate ranges, selectivity, or new functionalities; or to *de novo* engineer active sites from scratch.<sup>147</sup>

### 1.5.2 *Combination of rational design and directed evolution*

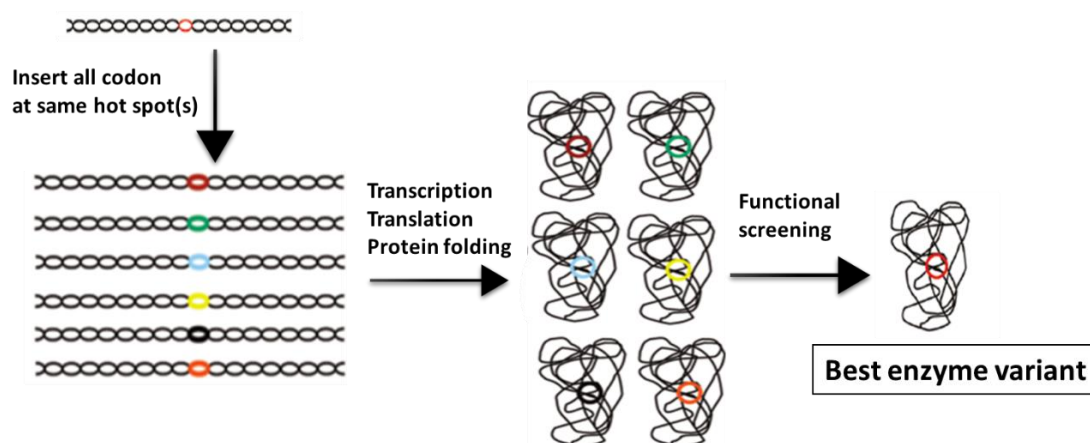
The directed evolution strategy explores enzymes residues to modify and stabilise a new system. Based on the “Darwin evolution” method, libraries of mutant enzymes are built by random mutagenesis to screen for an efficient and selective enzyme (**Figure 31, B**). Limits are encountered with this technique: tunable enzymes are required and a scaffold with a basic targeted catalytic activity is necessary. In addition, the randomness of this technique creates huge libraries of mutants that can overwhelm the experimental capacities in their preparation and screening. Introducing a new non-natural catalytic activity is a challenge with directed evolution method alone.<sup>148</sup>

Rational design is based on the knowledge of the protein structure, the catalytic reaction targeted, and any structure-function relationships (**Figure 31, A**). In the best case, the structure can be determined by X-ray crystallography or NMR spectroscopy, and additional kinetic experiments provide data on the mechanism of reaction. These data are then used to investigate promising binding sites or to spot specific amino acids of importance for anchoring and catalytic activity. By structural guided mutagenesis, rational design directly targets specific amino acids (from active or allosteric sites) responsible of an anchoring, an activity, or a structural modification. This technique allows to narrow the enzyme libraries toward a small number of more efficient mutants.<sup>149</sup> As an example, Wu and co-workers recently redesigned computationally an aspartase from *Bacillus* sp. YM55-1, introducing a non-natural activity and enlarging the substrate scope to build a small library of  $\beta$ -lyase. This enabled them to bypass multiple experimental trial and error.<sup>150</sup>

However, this methodology also shows some weaknesses. The rational design technique relies on prediction and knowledge of enzyme characteristics through its rigid crystal structure. It could thus fail to include the fine flexibility and interactions needed to achieve a high catalytic activity.

The latest studies in enzyme engineering tend to use a combination of both rational design and directed evolution for a complementary engineering of efficient artificial enzymes.

#### Site-saturation mutagenesis



**Figure 32:** Combination of Rational Design and Directed: studied enzyme structures and properties are used to select hot spot amino acids (rational design). These are replaced by all possible amino acids to create a smaller and efficient library (directed evolution). The library is screened, and the best variant is selected. Adapted from Gruber-Khadjawi and co-workers representation.<sup>147</sup>

#### 1.5.3 Application to AMs engineering

Impressive results have been achieved with rational design of artificial enzymes. Lu and co-workers extensively studied the repurposing of enzymes by structure knowledge design guided. They engineered *de novo* metal binding site by modification of the heme cavity of myoglobin.<sup>148,151</sup> Through site directed mutagenesis, the authors introduce a copper binding site and create an efficient artificial enzyme able to reduce O<sub>2</sub> into water.

In a similar way, Lin and co-workers modified the Mb to a dye-decolorizing peroxidase.<sup>152</sup> However, the resulting rationally designed AMs generally showed weak activity and selectivity. To accommodate the new function, the enzyme scaffold needs random mutations that are too difficult to predict. This is where directed evolution enters as an essential optimisation step.

The combined method of rational design / directed evolution has been extensively used. Baker and co-worker engineered an artificial organophosphate hydrolase from a native zinc adenosine deaminase.<sup>153</sup> Starting with computational modelling, they introduced mutations predicted by Rosetta software to create an initially low-active

AM. Based on this raw AM, random mutations increased the activity by around 2.500-fold. Similarly, Kim and co-workers changed the catalytic activity of an existing protein by inserting a  $\beta$ -lactamase activity inside a glyoxalase II.<sup>154</sup> They first extensively redesigned the active site with removal and insertion of entire loops. They then used directed evolution to fine-tune the new catalytic activity for a resulting AM which completely lost its original activity.

On a reverse strategy, Ward and co-workers investigated chemical and genetic optimisations to build a library of artificial imine reductases inside hCAII and screened for activity and selectivity. The best selected mutant ( $[(\eta^5\text{-Cp}^*)\text{Ir}(\text{pico})\text{Cl}] - \text{hCAII}$ ) crystal structure was then studied with Rosetta software. This software also used in previous examples, and extensively in other studies, is a computational package of macromolecular modelling, design and docking.<sup>155</sup> Ward and co-workers then applied site-directed mutagenesis on targeted residues to enhance activity and selectivity.<sup>84</sup>

These example highlights how rational design and directed evolution can be combined in a circle of optimisations.

This combined strategy was also applied to a more challenging AM engineering approach: the *de novo* design, or the creation of protein from scratch. Hilvert and co-worker developed a highly active and enantiospecific artificial zinc-dependent esterase.<sup>156</sup> Beginning with the *de novo* design of a homodimeric helix-turn-helix peptide, they built a four-helix-bundle protein with Zn(II). After several rounds of genetic optimisation, they obtained a mutant catalysing the hydrolysis of p-nitrophenyl acetate, with 10 000-fold activity increase. Based on their *de novo* protein, they recently predicted mutations through DFT calculation to create a diels-alderase by improving the substrate binding and transition state stabilisation. Several turns of directed evolution then improved the activity.<sup>157</sup>

#### 1.5.4 Computational design of AMs

##### 1.5.4.1 Computational methods

Computational tools can help in many ways in the redesign or the *de novo* construction of enzymes: from sequence homology to ligand binding analysis or



protein structural stability, computational expertise is essential for enzyme engineering.

As seen previously in the works done by Ward and co-workers, chemical and genetic optimisations are strategies of choice for the conception of AMs. However, this introduces a large amount of complexity in the computational design. The combination of cofactor, substrate and scaffold residues demand a lot of configuration and scaffold conformation exploration. Thus, computational techniques have evolved to use different powerful computational algorithms to take in consideration the function, the structure stability and the evolution of a protein.<sup>158</sup>

Starting with structurally unknown proteins, homology models can be built from high similarities with structurally characterised relatives.<sup>158</sup> Several algorithms know a high success in the protein structure prediction. For example, I-Tasser specialised in protein folding recognition or *de novo* modelling, and the software AlphaFold succeeded in the accurate prediction of protein fold and function from amino acid sequence alignment and comparison with proteins libraries.<sup>159,160</sup>

In AMs design, molecular modelling is the most applied technique. This approach regroups several methods to calculate the energies of a molecule in function of its geometry and its interactions with other molecules. The methods accuracy and approximations vary, where an accurate method needs high computational power, it will explore smaller space while for larger systems less accurate methods will be used. The challenge is then to select the right method by assessing the information needed and the technique cost.<sup>12</sup>

For the study of molecular interactions between a protein and its substrate /co-factor, docking is the most accessible and most prominently used approach (using molecular mechanic (MM) principles). The knowledge of the binding and release thermodynamics of these complexes uses more sophisticated methods, like free energy perturbation (FEP), for a calculation of the binding free energies, or the Molecular Dynamic (MD) simulations, used to understand the protein stability and flexibility. Finally, the catalytic mechanism is analysed with very accurate methods named quantum mechanics (QM) or a combination of QM/MM.

- ***Accurate methods: QM calculation***

Quantum mechanics (QM) methods describe accurately the electronic structure of a system, the chemical states of a reaction and the reaction mechanism. Two main families of QM methods have been applied to AM design: the Hartree-Fock (HF) and the density functional theory (DFT).<sup>12</sup>

The HF method determined the molecular electronic wave function of a system, built on the molecular orbitals. It approximates all interactions between electrons to a general wave function. Post-HF add electronic methods which present more accurate but also more resource consuming methods.

For DFT methodologies, the electronic calculations rely on the electron density. The approximate energy is calculated as a function of the electron density. DFT is an approximation method without any addition of time-consuming post-HF methods, and it can be used on a system with transition metal catalysts.

QM techniques are highly accurate to reproduce the electronic properties of a molecule. However, they are also highly time-consuming and computationally costly. QM methods are mainly used in small systems, for example a combination of the cofactor, the substrate, and few residues around.

- ***Medium methods: MM***

Molecular mechanics (MM) are used to model and describe large systems with reasonable accuracy, such as whole enzymes in water. This method relies on classical mechanics described by force fields. Unlike QM method, the electrons are not separated from the nuclei, and the molecules are represented by atoms and bonds. Force fields only describe two molecular systems by defining intra-molecular bonding terms (bonds, angles, dihedral angles) and intra- and inter-molecular non-bonding interactions (van der Waals and electrostatic interactions). MM methods represent much quicker and cheaper alternative, however, no electron transfer, and thus no chemical reactions can be described. These methods are compensating the limit of QM calculations: the ability to work on complete biological systems.

In biochemical interaction prediction, the simplest and most widely used MM technique is the computer-based docking screens. Docking is often utilised in drug design for the optimisation of lead ligands (molecules in complex with biomolecules) by prediction of stable ligand-protein complexes and the estimation of the binding affinity.

The docking methodology generates and evaluate ligand poses (orientation and conformation within the protein) to predict experimental binding modes and affinities. An algorithm evaluates the complementarity between a high-resolution structure and the molecules docked. Ligand poses are ranked through a score representing their stability and binding affinity.<sup>161</sup>

The docking methodology can be classified into three main classes that incorporate different levels of complexity to the protein and the ligand flexibility:<sup>162</sup>

- Rigid docking: within a rigid docking, both the enzyme and the ligand are kept rigid. Therefore, no alternative conformations are handled, and only translational and rotational freedom of the ligands are explored.
- Semi-flexible docking: several algorithms are used for the generation of ligand poses (e.g. systematic algorithms, stochastic or deterministic methods) which considers the conformational changes of the ligand besides the rotational and translational freedom. The protein is kept rigid.
- Flexible docking: the flexibility of the ligand is conserved while several parameters can set flexibility in the protein (e.g. residues side-chain and backbone flexibility, small overlaps between the protein and the ligand atoms).

To evaluate the binding affinity of the ligands poses, a scoring function is used. This adds favourable binding energies (e.g. hydrogen bonding, hydrophobic interactions) to a so-called binding score. Several methods of scoring function are available to predict the binding affinity (e.g. empirical scoring, force-field based, knowledge-based).<sup>163</sup>

Although the docking technique tries to reproduce the more accurately possible experimental results, some improvements have yet to be made, especially in the

binding affinity prediction. Basic scoring functions struggle to reflect experimental data and more sophisticated ones are too computationally expensive. Protein flexibility is also a challenging point. Most of the docking studies keep the structure rigid or slightly flexible, where, in nature, allosteric movement can impact the binding process.<sup>164</sup>

For a better understanding of the full ligand binding process, MD simulations are a more suitable method. Combined with FEP evaluations, the method gives higher accuracy for binding affinity of protein-ligand and protein-protein complexes.<sup>165</sup> In MD simulations, kinetic and potential energies are calculated, giving better insights than the static docking.<sup>12</sup> MD simulations are also used to analyse at atomic resolution the flexibility and stability of protein and ligand-protein complexes, however at a higher cost.<sup>166</sup>

#### - **Hybrid QM/MM**

At the interface of chemistry and biology, a hybrid QM/MM technique enables the in depth treatment of enzyme regions (binding sites) with a QM method coupled with an MM treatment of the entire protein structure.<sup>167</sup> This is an attractive strategy to accurately understand the details of a catalytic site and the reaction chemistry, with the treatment of the whole enzyme at a cheaper cost than with QM method alone.

Apart from mutant libraries creation or analysis of conventional enzymes, QM/MM is useful for the design of AMs. It is now frequent to combine QM/MM and docking or MD simulations for the exploration of specific binding properties, the prediction of energies or even the selection of suitable anchors.<sup>168</sup>

#### 1.5.4.2 Applications in artificial enzymes design

Computational methods have now been extensively used for enzyme engineering. For example, Kamerlin and co-workers used the empirical valence bond approach on the human diamine oxidase to understand its mechanism.<sup>169</sup> This method accurately investigates chemical reactivity, using MM representation in combination with QM electronic structure. This provides a fast and extensive knowledge of a chemical reactivity. Mulholland and co-workers also used QM/MM calculations to determine the mechanism of benzyl hydroxylation by the human cytochrome P450 2C9.<sup>170</sup>

Substrate unbinding kinetic was studied by Damborsky and co-workers using free energy calculation in their protein model haloalkane dehydrogenase.<sup>171</sup> All these information allow the identification of hit residues to modify for biocatalysis improvement.

The following section will focus on the use of computational tool for the engineering of AMs. As described before, the activity and selectivity of an AM are controlled by the binding of the metal catalyst and the substrate to the protein scaffold. Thus, the design of AMs involves modifications of the first coordination sphere (metal catalyst and residues in direct contact) and/or the second coordination sphere (enzyme scaffold). Another AM engineering approach is the *de novo* design, where new enzymes are entirely built from scratch. Computational modelling stands here as a critical tool to understand structural and biochemical processes, at low experimental cost.

#### **- *Design and optimisation of AMs***

A fundamental aspect of AMs engineering is the binding of the metal catalyst inside the scaffold. Molecular docking is often the first method of choice for the design of anchors binding to proteins, like in structure-based drug design.

Lu and co-worker used docking models for their initial AM to understand the covalent anchoring of their metal catalyst. They predicted the best fit inside the scaffold and then created a new AM with increased activity thanks to a tighter ligand affinity (dual anchoring).<sup>40</sup>

Protein-ligand docking studies have also been extensively used to understand experimental results. For example, Ricoux and co-workers ran docking experiments to explore their AM poor enantioselectivity (13 % ee).<sup>79</sup> Conclusions of the docking calculations identified a solvent exposed metal catalyst, due to a large occupancy of the cavity by the anchoring ligand. The substrate access to the metal was then less controlled and could explain the low enantioselectivity observed experimentally. Ward and co-workers also used docking to gain insight on the localisation of their metal complex catalyst.<sup>85</sup>

However, limitations to the docking approaches can be highlighted. Docking methods lack predictions of the dynamics of binding processes. They also poorly incorporate

the solvation and desolvation effects, which makes binding affinities prone to errors, particularly by neglecting entropic effects. Besides, the associated scoring functions use simple potential energy functions related to well parametrised force field, thus lacking contributions like polarisation effects and detailed estimation of proton affinity. Finally, transition metal catalysts are not yet routinely handled by docking. Standard molecular docking is not able to account for the change of electronic states and coordination of the metal or any conformational impact on residues during the binding process. In several studies the transition metal has been considered rigid, using a “dummy” atom with no impact on the scaffold upon binding.<sup>40,79,172</sup>

Refinement after protein-ligand docking with MD simulations or QM/MM calculations is then used for a more accurate process.<sup>173,174</sup>

A case protocol is described by Maréchal and co-workers.<sup>175</sup> The authors combined docking with QM/MM approaches for the introduction of Fe(Schiff base)-salophen inside a heme oxygenase. Their objective was to check out the changes induced upon binding of the complex during the activation process. First, based on the crystal structure, the authors optimised the Fe(Schiff base)-salophen using QM calculations. They proved that the experimental structure was indeed the resting state of the AM. They then docked the metal complex inside the haem cavity and studied the poses to identify all residues making contacts with the metal catalyst. Finally, to analyse the electronic transition and activation process, they generated all coordination possible of the complex residues-metal catalyst in the previously docked poses by refinement through QM/MM calculation. The results showed how the resting state (with a distorted metal complex binding) and later the activation process were driven by changes in the first coordination sphere.

A similar protocol was also followed by Ward and co-workers to analyse the catalytic mechanism of their first Sav artificial imine reductase and refine the structure.<sup>176</sup>

To further maximise enantioselectivity, computer modelling also facilitates investigations of the second coordination sphere upon metal complex-scaffold binding. Ward and co-workers, in collaboration with Baker group, tailored the second coordination sphere around a metal catalyst with the protein design software Rosetta Design.<sup>84</sup> To enhance the modest catalytic activity of their artificial imine reductase

$[(\eta^5\text{-Cp}^*)\text{Ir}(\text{pico})\text{Cl}]$  - WT hCAII), they selected and mutated specific amino acid residues in the protein-ligand binding interface. Their goal was to firmly anchor the metal catalyst in one location that favours the interactions with the substrate. Based on the AM crystal structure (PDB 3ZP9), four mutants were developed to improve affinity between metal catalyst and protein. Later characterized experimentally, the designed mutants showed increased affinity to a 50-fold, activity to a 10-fold and enantioselectivity to 96 % ee-(*S*) salsolidine.

The other essential aspect of AMs design is the substrate binding to the enzyme. Molecular docking is again a gold standard to understand the interactions upon binding. Coupled with QM techniques, docking also helps to understand the catalytic pathway and any interactions between the substrate and the metal catalyst.

Ricoux and co-worker engineered a Mn(III)-*meso*-tetrakis(*p*-carboxyphenyl) porphyrin inserted within a xylanase. They analysed through docking the epoxidation of aromatic alkenes substrate.<sup>177</sup> The first molecular docking run was to identify the characteristics of the metal complex binding to the xylanase. Several aromatic substrates were subsequently docked inside the previous best docking pose. Adequate new poses were selected with proximity between the metal and the alkene double bond. Some residues in the active site pocket were spotted as regulators of the substrate access to the active site and the proximity with the metal catalyst. These results were later confirmed by experimental tests.

Furthermore, computational modelling has also been used for the identification of novel substrates for AM. An example is the work of Ménage and co-workers on their artificial monooxygenase.<sup>178</sup> The authors developed four AMs from a transport protein NikA, using supramolecular anchoring of several iron metal complexes. They then docked a set of sulfides substrates to identify catalytic oxidation of their hybrids. The molecular docking concept was then proven by experimental tests on the selected substrates.

#### - ***De novo design of artificial enzymes***

Computational modelling tools are already used in combination with directed evolution to either design catalyst-scaffold-substrate bindings or increase the activity

and selectivity of AMs. However, one domain is still challenging for computational modelling: the *de novo* design of AMs from scratch.

A typical protocol starts with defining a catalytic reaction and its mechanism. Then, a “theozyme” is modelled quantum mechanically: a minimum three-dimensional model of an active site structure is defined, composed of a cofactor, a substrate and the amino acids surrounding. The QM method calculate the optimal arrangement of the active site to stabilise the transition state of the targeted reaction. Using explorative algorithm, this model is then compared to existing protein in order to define a scaffold.<sup>179,180</sup>

Some first approaches have already been developed. Baker and co-workers used the most common Rosetta software pack to design an enzyme model for Diels-Alderase reaction.<sup>181</sup> RosettaMatch was first used to find a corresponding scaffold to the theozyme and RosettaDesign followed to optimise the transition state structure of each matches. The rigidity of the predicted models differs from the conformation changes of a scaffold during catalytic process, and a refinement is generally executed with QM/MM calculations.

Other studies have followed a similar protocol for different reactions. For example, Bolon and Mayo worked on hydrolysis reactions,<sup>182</sup> Houk and co-workers on Kemp elimination<sup>183</sup> or Baker and co-worker on Morita-Baylis-Hillman reaction.<sup>184</sup> Clearly, the computational strategy alone is not able to produce a sustainable level of activity for the new artificial enzymes compare to other catalyst types. Further studies from these groups combined subsequent computational optimizations and directed evolution, to create hybrids with enhanced effectiveness.<sup>153,185,186</sup> These results are still challenged due to the simplicity of the catalytic reactions targeted, and thus the luck could also be considered in the success of such strategy.<sup>12</sup>

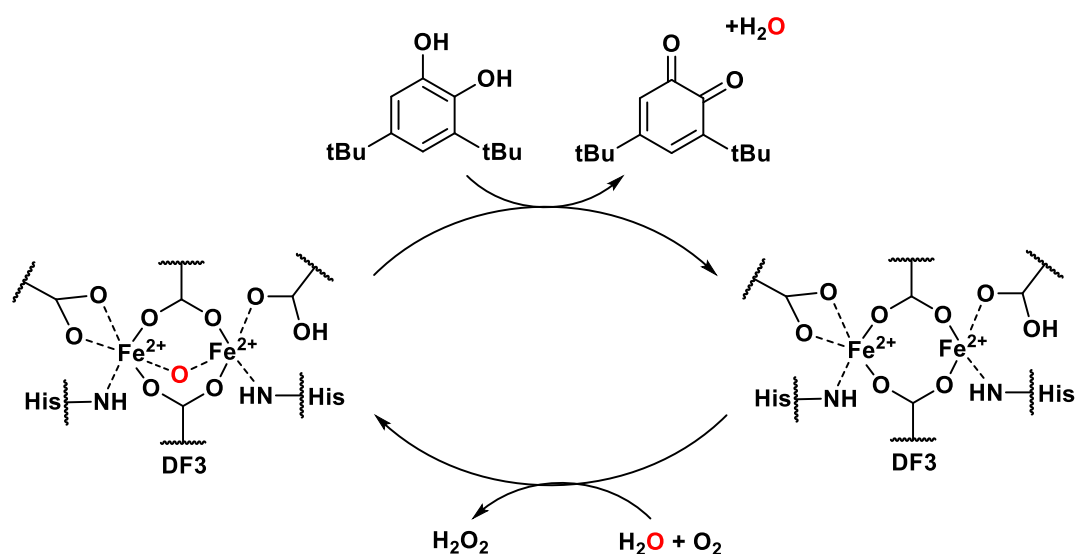
When considering artificial metalloenzymes, the complexity of a *de novo* design is at a higher degree. As seen before, activity is regulated by the binding of the metal catalyst. Thus, the theozyme models would have to also account for the correct coordination of the metal inside the scaffold: first and second coordination spheres must be designed.

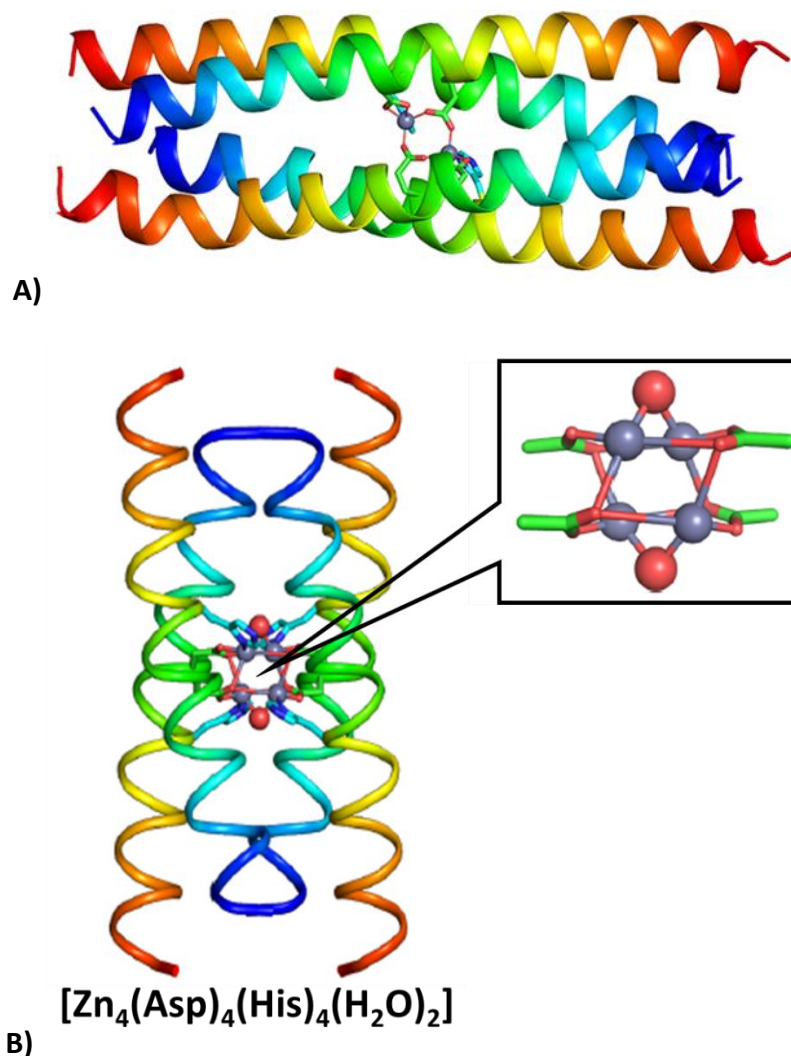


To date, no efficient *de novo* AM has been engineered from computational modelling, but enormous progresses have been achieved toward this objective.

On the second coordination sphere design, models of protein folding were developed by Wolynes and co-workers<sup>187</sup> and Baker and co-workers.<sup>188</sup> Experimental results confirmed the predictions of the new protein folding. The design of the first coordination sphere has also been studied. Hellinga and co-workers<sup>184</sup> achieved the introduction of a novel iron binding site inside an *E. coli* thioredoxin and Baker and co-workers introduce unnatural amino acids to engineer a new binding site.<sup>25</sup>

Finally, the first *de novo* proto-enzymes have been reported based on a rough starting point, the *Due Ferri* (di-iron) helical bundle created by DeGrado and co-workers with mathematical equations.<sup>189</sup> A di-iron oxidase, an oxygenase and a ferroxidase were engineered by DeGrado group using computational protein design (**Figure 33, A**).<sup>189,190,191</sup> By modelling the protein folding in concert with the metal binding site, the authors succeeded to insert within the four-helix bundle scaffold a transition metal cluster composed of four  $\text{Zn}^{2+}$  and 16 polar residue side chains (carboxylates oxygens, imidazole), all connected by hydrogen bonds (PS1, **Figure 33, B**).<sup>192</sup> The process used automated sequence design and Rosetta methodology on the *Due Ferri* starting point to develop the final proto-enzyme, followed by QM/MM methods to understand the engineered structures.





**Figure 33:** DeGrado *de novo* proto-enzymes. **A)** Top: Proposed mechanism of the catechol oxidation reaction catalysed by DF3. Bottom: Representation of a four chain DF structure (similar to DF3). **B)** X-ray structure of PS1 (PDB 5WLM) with a close view of the zinc cluster. Figures adapted from DeGrado and co-workers.<sup>189</sup>

### 1.5.5 Conclusion

Computational tools must be considered an important strategy for AMs engineering, in combination with experimental work. So far, *in silico* studies have proven to be useful for a large range of information (from coordinated interactions to binding process mechanism). The methods are in constant evolution to be more fine-tuned and more powerful, for the accurate prediction of all aspects of the biochemical transformations during a catalytic reaction.

## 2 Aim and objectives

### 2.1 Research overall aim

Protein engineering is now an advanced biotechnological strategy. However, the introduction of non-native activity is still challenging. Artificial metalloenzymes (AMs) are becoming a powerful approach to address this question. The introduction of a metal catalyst in an enzyme scaffold leads to a renewable catalyst, able to catalyse green reactions efficiently and selectively.

In the AM field, the supramolecular anchoring achieved a great success with the biotin-(strep)avidin and LmrR receptor technologies. However, no relevant evidence for other efficient supramolecular systems have been privileged yet. In addition, the use of enzymatic scaffolds has been overlooked compared to their non-enzymatic counterparts. The existing architecture of the enzymatic site within an enzymatic chassis can be advantageous. For example, the natural evolution of the pockets provides a proximity between cofactor and substrate, which can be exploited in the non-covalent binding of a metal catalyst.

In previous works, the design of AMs has been supported by computational modelling, using predominantly docking to help in the optimisation of AMs. Computational docking is rarely used to design the initial AMs from the beginning, but more often to optimise and understand AM mechanism. Starting the engineering of AMs with computational modelling would be a gain of time and experimental resources.

The aim of this project is to computationally design anchors for the binding of metal catalysts inside a cofactor pocket, as starting point to engineer new AMs. The focus of this work is to expand the enzyme functionality of alcohol dehydrogenases (ADHs) from reduction of carbonyls to the asymmetric reduction of imines, with high enantioselectivity.

A literature search concluded that there was interest in imine reduction to chiral amine. These are important pharmacophores found as building blocks in many active molecules with applications in the chemical and pharmaceutical industries. The use of a  $d^6$  piano-stool transition metal was thus prioritised since such compounds are

highly active in imine reduction. The choice of ADHs as an enzymatic scaffold, for the reduction of imine, is based on the structural similarity of their natural ketone substrate C=O and the imine structures C=N. Furthermore, the natural evolution of the catalytic pockets in ADHs contributes to the strong binding affinity to their natural NAD(P)H cofactor, along with high stability and proximity to the substrate. This work sets out to create analogues of NAD(P)H as anchors for the transition metal catalyst. Two ADHs enzymes were selected for the AM scaffolds: the Horse Liver Alcohol Dehydrogenase (HLADH) and the *Thermoanaerobacter brockii* Alcohol Dehydrogenase (TbADH).

## 2.2 Research objectives

### 2.2.1 *In silico design of supramolecular anchors for transition metal catalysts*

The initial focus of this project is the computational design of appropriate small anchor molecules, with high affinity for the NAD(P)H binding pocket of the two selected ADHs. Docking methods will be set up to first assess the relative affinity of natural cofactors, mimics thereof and large libraries of existing compounds, within HLADH and TbADH. The objective is to gain detailed knowledge of the binding pocket and the non-covalent interactions essentials to ensure strong affinity of small molecules within ADHs. These conclusions will lead to the creation of a novel anchor structure model, with the required functionalities to enable the binding of a metal catalyst. A final computational modelling protocol will investigate the anchor binding and evaluate the impact of metalation on the binding orientation inside the pocket of ADHs. The resulting docking will provide a refined ranked list of molecules with high affinity and appropriate location.

### 2.2.2 *Chemical synthesis of the designed transition metal catalysts*

This objective includes the design of a short synthetic pathway and the creation of a library of small anchors. Based on the results predicted by the docking protocol, the anchors which demonstrated the highest affinity for ADHs will be selected for synthesis. All synthesised compounds will be metalated with iridium to give the

transition metal catalysts. Their catalytic reactivity will then be tested to ascertain their effectiveness in the transfer hydrogenation of imines.

### *2.2.3 Experimental creation and functionality testing of a designed AM*

The final part of the project concentrates in the construction of AMs. To do so, a reliable method for the expression and purification of TbADH, HLADH and also a TbADH mutant which lacks the catalytic zinc ion will be used. The previously synthesised transition metal complexes will be non-covalently assembled inside the purified ADHs. The high affinity of anchors binding will be assessed by isothermal titration calorimetry and inhibition assays. The catalytic activity and selectivity of the resulting AM will be evaluated by HPLC analysis and will be compared to the free catalyst complexes. This will allow a critique of the computational results and therefore assessment of its capacity as a predictive tool to design AMs.

### 3 Docking guided design of artificial metal catalysts with affinity for alcohol dehydrogenases.

#### 3.1 Introduction

In the following chapter, the computational methods used throughout the *in silico* study are detailed. The chapter focuses on computational applications and their theoretical background without accounting for hardware considerations. The process toward the selection of molecules as metal anchors with high affinity for both HLADH and TbADH is described.

The computational methods were developed to understand key features of the cofactor binding site of these enzymes and identify the strongest natural non-covalent interactions, to promote the binding of synthetic molecules. Co-crystallized structures of HLADH and TbADH were selected for docking studies, based on previous crystal structures of purified enzymes in complex with their cofactor or substrate (respectively PDB 4XD2<sup>193</sup> and PDB 1YKF<sup>97</sup>).

#### 3.2 Computational methods

##### 3.2.1 *Molecular Docking*

For the design of molecular docking protocols, the Schrodinger suite of programs and the Maestro graphical interphase were used. The theoretical background for molecular docking was presented in the section 1.5.4.1, page 43.

##### 3.2.1.1 Ligand preparation

All compounds included in this work were prepared using Ligprep from Schrödinger suite,<sup>194</sup> with OPLS3 force field.<sup>195</sup> Generation of all possible protonation and ionisation states combinations was performed for each ligand by using Epik<sup>196</sup> in aqueous solution at pH of 7.0 +/- 2.0. Whenever the ligand contained a ribose, the same stereochemistry as of the natural cofactor was kept.

##### 3.2.1.2 Protein preparation

The crystal structures of the enzymes were obtained from the Protein Data Bank (HLADH PDB 4XD2<sup>193</sup>, HLADH with substrate PDB 4DWV<sup>197</sup> and TbADH PDB 1YKF<sup>97</sup>)

and prepared using the Maestro Protein Preparation Wizard in the Schrödinger suite.<sup>198</sup> The structure integrity was analysed and missing hydrogen atoms and residues side chains were added to the structure by Prime-refinement<sup>199</sup> throughout the pre-processing. During the refinement, water molecules with less than three hydrogen bonds to other atoms were removed, which resulted in no more water in the binding site. The selection of the position of hydroxyl and thiol hydrogens, the protonation/tautomer states and the “flip” assignment of aspartic acid, glutamic acid, arginine, lysine and histidine were adjusted at pH = 7.0 using PROPKA.<sup>200</sup> Finally, the structures were minimized using the OPLS3 force field with a RMSD = 0.3 Å maximal displacement of non-hydrogen atoms as convergence criterion.

The proteins were additionally prepared without the natural zinc ion in the active site, by removing it from the crystal structure using the Protein Preparation Wizard. Furthermore, several amino acids present in the active site were mutated to alanine: H67, C46 and C174 in HLADH and C37, H59, E60 and D150 in TbADH. These mutations allowed to lose the zinc ion and to make space at the catalytic site.

#### 3.2.1.3 Grid preparation

In order to define the pocket of interest for the docking, a grid was set around the defined docking site. At the beginning of this work, the same grid preparation setup was chosen for all docking methods. The enclosure box was defined from the optimized protein structure at the centroid of the active site (10 Å radius around the co-crystallized ligand). No constraints were added. The standard settings of van der Waals scaling factor of 1.0 for nonpolar atoms was conserved. Nonpolar atoms were defined with absolute value of partial atomic charges  $\leq 0.25$ .

#### 3.2.1.4 Glide docking

Prepared ligand conformers were flexibly docked into the receptor grid using Glide SP and XP procedures.<sup>201,202,203</sup> In this method, the receptor was kept rigid.

Glide docking uses a technique of exhaustive systematic search to predict orientation, conformation, and binding position of a ligand inside a receptor pocket. The following steps are undertaken during a Glide docking procedure:

- 1) Collection of ligand conformations (from exhaustive enumeration of ligand torsions)
- 2) Initial screening inside the predefined grid space. This results in the selection of ligand poses (“poses” refers to the orientation, position, core conformation and rotamer-group conformation of the ligand inside the receptor).
- 3) Refinement of pre-selected poses by analysis of ligand torsions in the field of the receptor.
- 4) Minimization of ligand poses with full flexibility. A Monte Carlo procedure analyses nearby torsional minima and selects the minimum energy poses.
- 5) Final scoring applied to the poses predict binding affinity (GlideScore).

- ***Glide standard precision docking (Glide SP)***

This is the classical method which uses exhaustive sampling method.<sup>201,202</sup> The default parameters as defined in the Schrodinger program were kept, OPLS3<sup>195</sup> was used as force field and the variation of conformation was applied.

The four best docked structures were saved for each ligand, based on the Emodel score which allows a better discrimination between the conformers of a ligand (it combines the non-bonded interaction energy and the excess internal energy of the generated ligand conformation). Subsequently, the Glide score was the focus outcome of the Glide SP docking. Contrarily to Emodel score, the Glide score ranks the different ligands and maximizes the discrimination between ligands that bind strongly inside the pocket to the ones that have little affinity. Glide Score is an empirical scoring, taking in account the physics of the binding process (e.g. lipophilic-lipophilic, hydrogen bond, protein-ligand coulomb-vdW energies) and the displacement of water molecules by a ligand. It is an estimation of the ligand free energy binding (the more negative, the better).

The Emodel score was used to select the best ligand conformer and then the Glide score ranked all these ligands structure for a determination of the best binder.

- ***Glide extra precision docking (Glide XP)***

Glide XP performs a more extended sampling methodology and has an optimized scoring function. It penalizes some false positives, which Glide SP does not eliminate,



by adding new elements in the scoring: desolvation penalties applied to polar and charged groups and recognition of groups that increase affinity.<sup>203</sup>

Glide XP is more time consuming than the SP version. Therefore, this method is better when used on predefined ligand poses, after a pre-screening by Glide SP. The Glide XP score is more sophisticated by identifying better complementarities between the ligand and the protein (penalties are assigned to expected unfavourable energy complexes).

#### 3.2.1.5 Refinement of the ranking: MM-GBSA method

The MM-GBSA method is a molecular mechanic energies method, coupled with generalized Born model surface area continuum solvation. This end point method is based on the end states difference of the free ligand, the free receptor, and the complex. It is an effective method with a satisfying compromise between a good re-ranking of Glide XP docking poses and a lower time consumption, and is used to give an estimation of the binding affinity between a ligand and its receptor in a complex.<sup>204</sup> The best ten XP scored poses were retained for MM-GBSA predictions.

The relative free binding energy ( $\Delta G$ ) was calculated using Prime/MM-GBSA software from the Schrodinger Suite, according to the following equation:

$$\Delta G = E_C - (E_L + E_R) \quad (1)$$

Where  $E_C$ ,  $E_L$  and  $E_R$  represent the free energy of a state (complex, ligand, and receptor respectively). The software uses different energy properties (structure, binding, and strain) to obtain total energies applied to the different states:

$$E = E_{MM} + G_{solv} - TS \quad (2)$$

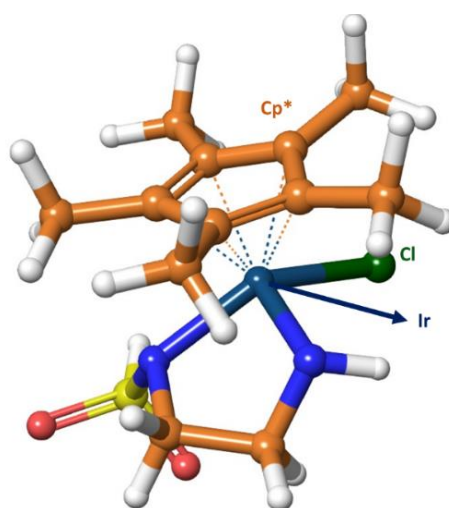
Where  $E_{MM}$  represents the molecular mechanics energy,  $G_{solv}$  the solvation free energy and  $S$  the entropy of the system. The contribution of various factors is also considered (e.g. H-bond, vdW, solvation energy) in this protocol.<sup>205</sup> The VSGB solvation model (Surface Generalized Born Model and Variable Dielectric) was selected to approximate the solvation free energy and correction of interactions (H-bond, hydrophobic,  $\pi$ - $\pi$ ).<sup>206</sup> The OPLS3 force field was applied and flexibility was added to residues at 5.0 Å from the ligand position in the binding pocket. A Prime

minimisation was used on the residues selected in the flexible region, without any constraints.

#### 3.2.1.6 Metal complex docking

For docking of metal, a challenge arises from the software being unable to handle ligand structures incorporating complexed metal ions. In the Schrodinger software, force field parameters available in Glide are designed to treat structures with zinc, copper, or iron only. However, this project deals with the treatment of transition metals such as iridium (Ir), rhodium (Rh) or ruthenium (Ru). To dock structures containing transition metal, the settings of the OPLS3 force field were modified. In this work, the experiments were attempted with metal complexes made of Ir.

At first, structures containing Ir metal were created by using the Maestro interface to build a pyramidal metal centre, substituted by a pentamethylcyclopentadienyl moiety (Cp\*, **Figure 34**).



**Figure 34:** Iridium coordinated with chloride, bidentate sulphonamide ligand and Cp\*. This transition metal complex was built using the Maestro building panel.

Zero bonds order (no electron bonding) were applied to the Cp\*metal ion ligand (dash lines, **Figure 34**). Force fields have difficulties to characterise metal coordination and there is no criterion for covalent bonds between metal and ligands. Thus, zero bond order connections represented the metal coordination geometry. The different metal ligands were left flexible.

#### - ***Grid preparation***

In the chapter section involving the transition metal complex docking, the enclosure box was defined from the optimized mutated protein structures (HLADH H67A, C46A, C174A and TbADH C37A, H59A, E60A and D150A), by a selection of amino acid residues defining the centre of the box to be the area of the NADH nicotinamide ring. No constraints were added. Van der Waals scaling factor of 1.0 and partial charge  $\leq 0.25$  were conserved. These grid modifications allowed to focus the binding of metal complexes around the catalysis area.

#### 3.2.2 *Induced Fit Docking (IFD) calculations*

The IFD method considers the receptor pocket to be flexible around the ligands docked.<sup>207</sup> It combines Glide,<sup>202</sup> to obtain the possible binding poses, and Prime<sup>199</sup> which associates conformational changes of the receptor structure (refinement of active site residues).<sup>208</sup> The following steps were undertaken:

- 1) Docking of the ligand with Glide
- 2) Prime refinement to reorient amino acid side chains for each ligand pose found by Glide.
- 3) Minimisation of residues and the ligand.
- 4) Re-docking of the ligand in the receptor active site, using Glide, and minimisation and ranking of complex poses by IFDscore.

The IFD parameters were defined as follows: Prime refine residues up to 5.0 Å to the ligand position. The first docking used Glide SP and subsequent re-docking steps used Glide XP for refinement, which gave a final Glide XP score for each ligand structures.

### 3.3 Results and discussion

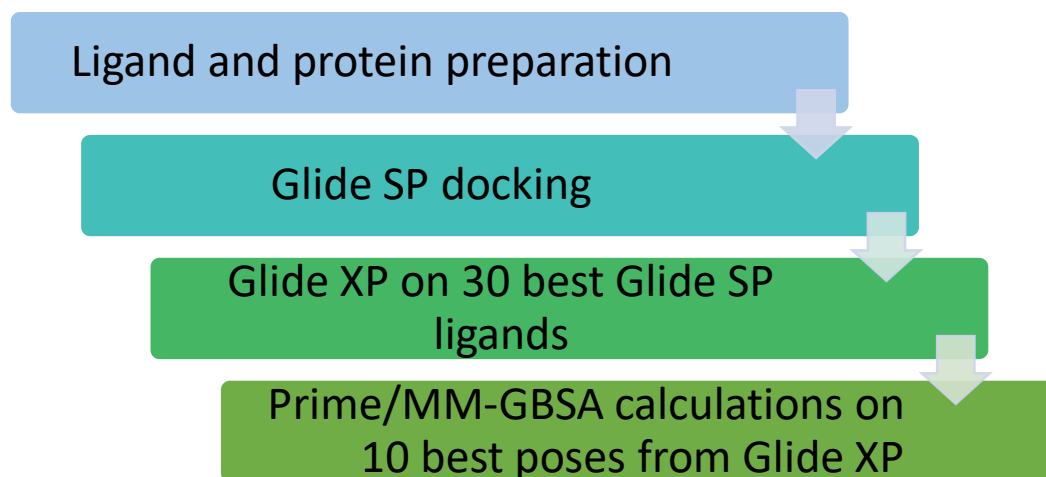
The most employed method to predict biomolecular interactions between a protein and a molecule is the docking.<sup>12</sup> By providing binding structures and binding scores, docking tools give an insight into the pocket geometry, the binding conformation, the chemical interactions and an approximation of the affinity.

However, precautions are to be taken in the interpretation of docking results, especially in the catalytic domain. The approximation resides mainly on the scoring function, which lack certainty in the representation of binding affinity.

### 3.3.1 Systematic docking evaluations to study the cofactor binding pocket

#### 3.3.1.1 Design and validation of a docking protocol

Docking studies are a promising tool to predict ligands configurations and their affinity to a protein binding site. A docking protocol was prepared to filter the ligands that bind with high affinity to the enzymes out of large sets of structures and bound conformations:



**Figure 35:** Established docking protocol used to filter various ligands in order to select lead structures. The two Glide dockings give a rough ranking of structures, followed by a refinement using MM-GBSA method.

The protocol started with a raw filtering of the NAD(P)H, fragments of the cofactors and libraries of ligands structures inside HLADH and TbADH cofactor binding site. First, the Glide SP method was applied, followed by Glide XP (elimination of false positives). This rough ranking of structures separated the binding ligands from the less or non-binding ones. However, this score had a too low accuracy for a hierarchical separation inside the group of binding ligands and for the determination of a potential lead compound. Estimation of the free binding energy of enzyme-ligand complexes would provide more information on the affinity for a ligand, the stability of a system and a more precise ranking. As an example, molecular dynamic (MD) simulations combined with free energy perturbation (FEP) are a powerful tool to retain accurate binding affinities. This strategy is more accurate toward experimental data than Glide scores. However, MD simulations are also time expensive, and thus difficult to use on a vast number of ligands.<sup>209</sup>

An interesting alternative was the Molecular Mechanics Generalized Born and Surface Area (MM-GBSA) method, for the estimation of binding free energies. The

resulting final scores predicted the relative binding affinities and provided a ranking between selected binding ligands.

The quality and efficiency of this protocol was tested by re-docking the co-crystallized NAD(H) cofactor inside the protein and comparing the docking poses through calculations of root mean square deviation (RMSD) by using the superposition panel in Maestro. A good docking pose is highlighted by an RMSD value  $< 2.0$  Å from the co-crystallised cofactor (**Table 1**).<sup>210</sup>

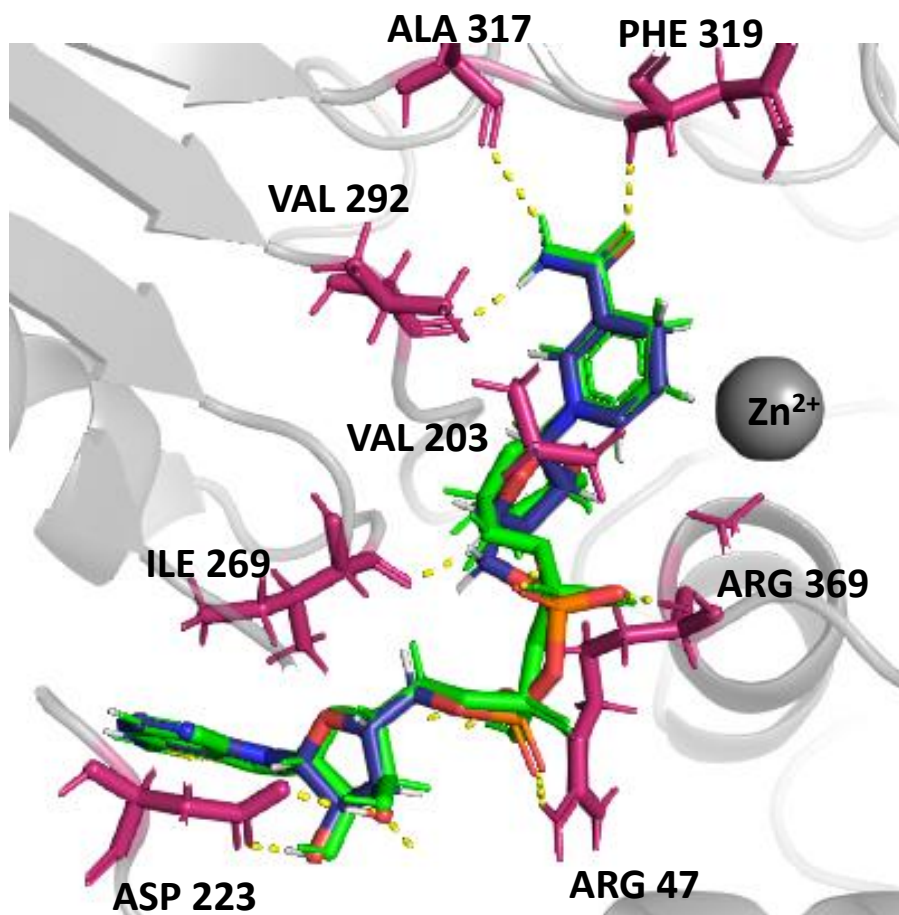
Two HLADH structures were used in which the enzyme was crystallised with only the cofactor NADH (PDB 4XD2<sup>211</sup>) and with the cofactor NAD<sup>+</sup> plus an alcohol substrate (2, 3, 4, 5, 6-pentafluorobenzyl alcohol, PDB 4DWV<sup>197</sup>). In each crystal structure type, another round of docking was conducted without the catalytic zinc ion to evaluate the impact of Zn<sup>2+</sup> removal on the orientation and binding affinity of the ligands. This gave more information on whether binding space can be increased in the absence of the Zn<sup>2+</sup>, to allow its replacement by a different transition metal complex. The results are presented in **Table 1** below:

**Table 1:** Cofactors docking scores represented by Glide XP (kcal mol<sup>-1</sup>) and MM-GBSA calculations ( $\Delta G$  in kcal mol<sup>-1</sup>). The RMSD values compare the dock poses to the co-crystallised cofactors in the X-ray structures.

Enzyme	PDB	Ligand	Resolution (Å)	Glide XP Score	RMSD (Å) on Glide XP pose	MM-GBSA	RMSD (Å) on MM-GBSA pose
HLADH	4XD2 <sup>211</sup>	NADH	1.1	-14.8	0.32	-80.6	0.33
HLADH no Zn <sup>2+</sup>				-15.1	0.41	-83.6	0.37
HLADH + substrate	4DWV <sup>197</sup>	NAD <sup>+</sup>	1.14	-11.7	0.61	-101.2	0.98
HLADH + substrate no Zn <sup>2+</sup>				-13.1	0.49	-124.6	0.97

For all HLADH structures, the superposition displayed a RMSD values  $< 2.0$  Å, attesting accurate docking poses. The correct re-docking of NADH/NAD<sup>+</sup> gave confirmation of an efficient docking protocol for HLADH *in silico* study, which was able to relocate the

cofactor in its original position and conformation. The superimposition of the docked structure and the crystal structure of NADH inside HLADH is shown in **Figure 36**.



**Figure 36:** NADH (blue) re-docked inside HLADH and superposed to the co-crystallised cofactor (green). Hydrogen bonding contacts are represented in dashed lines and the amino acids in contact with the ligands are labelled and shown in pink.

The removal of  $\text{Zn}^{2+}$  led to no variation of values or orientation inside the binding site, and only a slight improvement in the binding free energy was seen with MM-GBSA calculations. These first results implied little impact of  $\text{Zn}^{2+}$  in the binding of cofactors. As suggested in the literature,  $\text{Zn}^{2+}$  is used to act as Lewis acid and activate the carbonyl for hydride transfer.<sup>91</sup> Thus, based on literature and first docking conclusions, the zinc role is mainly of catalytic nature and may have no influence on the cofactor binding.

In summary, the designed docking protocol was proven accurate for the analysis of ligands inside the different HLADH structures. The following results were obtained using this protocol on various ligand structures inside the two HLADHs crystal structures.

### 3.3.1.2 Docking into HLADH without substrate (PDB 4XD2)

After validation of the protocol, docking of other molecule structures was performed and compared to co-crystallized NAD(H) cofactors. The experiments started with fragments of NAD(P)H cofactors, followed by NADH mimics from the literature and previously designed compounds in the laboratory, to end with a screening of large compound libraries. The aim of this methodology was to understand key features of the cofactor binding pocket and to discriminate the influence of different fragments from the relatively big cofactor structure. The diversity of structures explored within the screening of large compound libraries further helped to better understand the potential of the cofactor binding pocket environment to host alternative molecules with different functionalities.

### 3.3.1.3 Docking of NAD(P)H and NAD(P)<sup>+</sup> cofactors

The various natural cofactors in reduced and oxidised states (NADH, NAD<sup>+</sup>, NADP<sup>+</sup> and NADPH) were compared for the evaluation of differences in their binding affinities to HLADH.

**Table 2:** Cofactor docking inside HLADH (PDB 4XD2). Glide XP score (kcal mol<sup>-1</sup>) and the free binding energy  $\Delta G$  (kcal mol<sup>-1</sup>) from MM-GBSA calculations are presented for all cofactors. The lower the score, the better the binding affinity.

Enzyme	NADH		NAD <sup>+</sup>		NADPH		NADP <sup>+</sup>	
	XP Glide	$\Delta G$	XP Glide	$\Delta G$	XP Glide	$\Delta G$	XP Glide	$\Delta G$
HLADH	-14.8	-80.6	-11.5	-92.7	-14.9	-73.5	-11.1	-83.4
HLADH no Zn <sup>2+</sup>	-15.1	-83.6	-12.0	-125.7	-13.2	-70.3	-11.8	-99.1

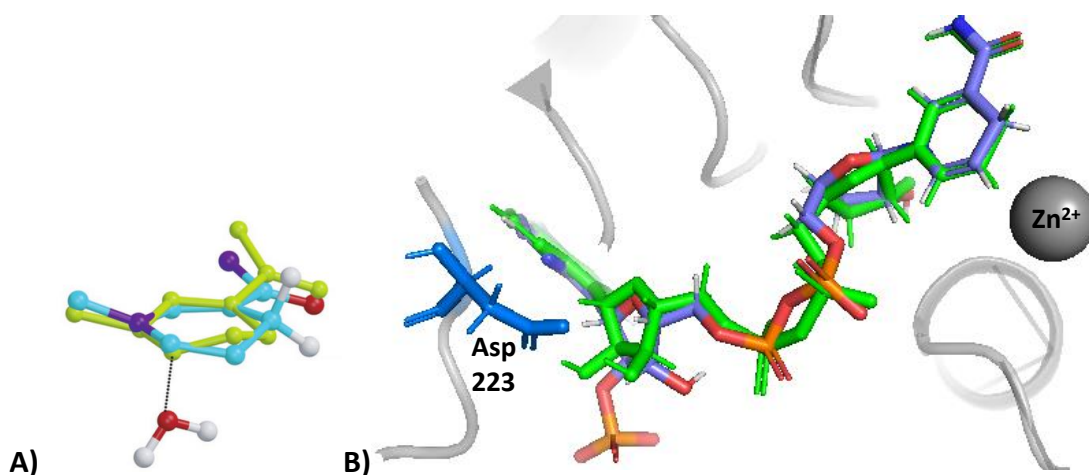
The comparison of these results led to a more complex picture of binding. According to both scores, the binding of cofactors NADH and NADPH were of similar order, with a slight preference for NAD(H) binding. A minor binding difference was observed with Glide XP score between the oxidised and the reduced forms of the cofactors. Then, in the MM-GBSA calculation, the ranking established a significantly better affinity for complexes with oxidized forms (NAD<sup>+</sup>, NADP<sup>+</sup>).

All NAD(H) nicotinamide ring structures appeared flat inside the pocket of HLADH. However, the QM calculation works of Meijers and co-workers demonstrated a distortion of the pyridine ring in NADH, later followed by Rubach and Plapp

calculations, suggesting a boat like conformation during the transition state (**Figure 37A**).<sup>212,213</sup> The docking results might indicate that the force field parameters were not accurate enough to describe the difference between the oxidised and reduced structures (difference in solvation free energy of the cation), which in turn might lead to inaccuracies of the docking scores in comparison. Thus, the comparison between docking of reduced and oxidised cofactors needs to be treated with caution at this stage.

To the best of our knowledge, there is no direct information available in the literature about binding affinity studies to HLADH of the two cofactors NADH and NADPH. Examining the work detailed in literature (section 1.3.2.2, page 24), a specificity of HLADH to NADH was estimated, but it is unclear whether NADPH does not bind inside HLADH or bind in an inactive conformation. The literature still suggests the binding of NADPH is hindered by a repulsion between the ribose phosphate and D223.<sup>214,215</sup>

A tendency to bind NADH better than NADPH was also found in both docking approaches. From the visual analysis, a slight difference was showed between NADPH and NADH position and contacts, with the adenosine ribose phosphate group of NADPH orienting the ribose away from D223 (**Figure 37B**).



**Figure 37:** **A)** Superposition of the crystal nicotinamide ring of NADH (yellow) and the structure from QM calculations (cyan) from Meijers and co-workers **B)** Superposition of NADPH cofactor (blue) with the NADH (green) from docking experiment. The highlighted D223 repulsed the NADPH ribose phosphate from the adenine group. By contrast, the ribose of NADH made contacts with D223.

In general, all the cofactors bound with significant high affinity values, predicting a relative high affinity to HLADH. These results gave a reference for further studies of different ligands structures. However, more experiments will be necessary to inform

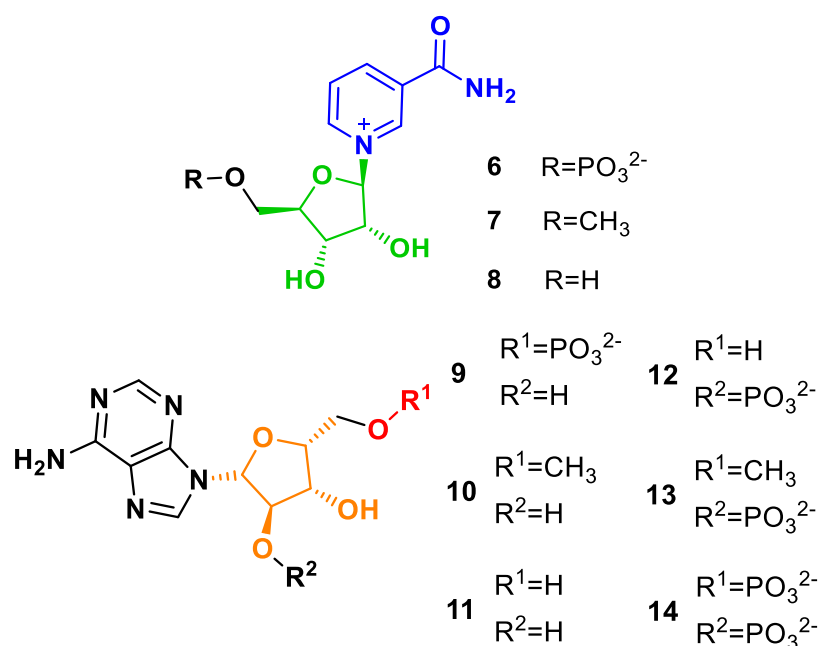


if the binding affinity ranking between the different cofactors in the docking simulations is of quality.

#### 3.3.1.4 Docking of NAD(P)H fragments

From the literature, it is unclear whether strong NADH binding is due to the adenosine fragment, the nicotinamide part or both. The screening of molecules fragments is helpful in order to better understand key binding interactions. As the molecules are smaller, they explore a larger part of the binding site. These smaller structures are likely to form few bonds with the target binding site, but these links have to be strong to form a stable complex.<sup>216</sup> Thus, this fragment screening was applied to obtain a better idea on which NADH part is responsible for the affinity inside HLADH.

The NAD(P)H cofactor was split into two fragments between the two phosphate groups (**Figure 38**): the nicotinamide riboside structures bearing modifications at the ribose primary hydroxyl group (fragments **6** to **8**) and the adenosine fragments with or without the corresponding ribose monophosphate (fragments **9**, **10**, **11** and **12**, **13** and **14**, respectively).



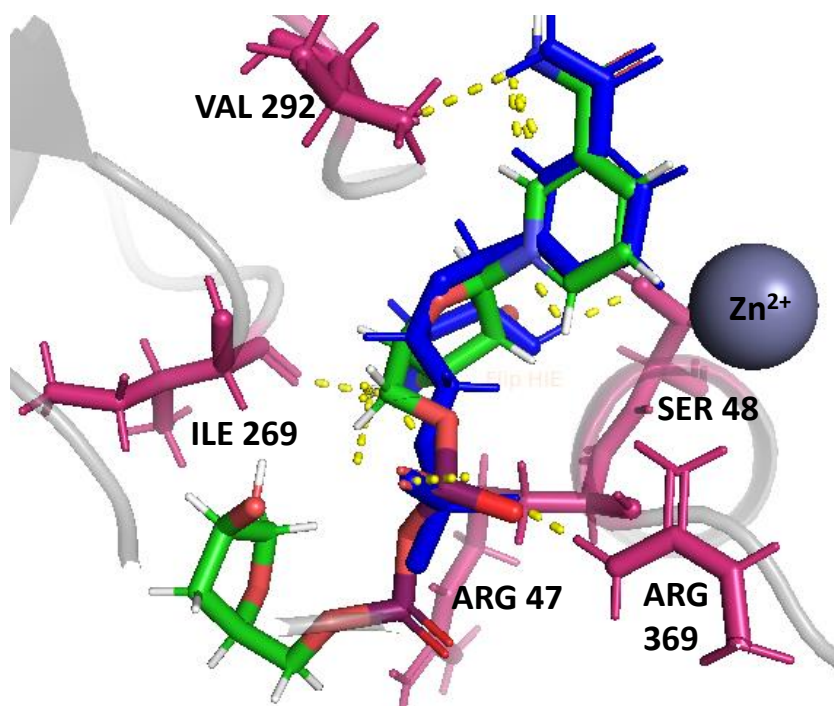
**Figure 38:** Fragments of NAD(P)H structure based on the nicotinamide and adenosine parts, designed for the docking inside HLADH.

**Table 3:** Glide XP (kcal mol<sup>-1</sup>) and MM-GBSA ( $\Delta G$  in kcal mol<sup>-1</sup>) results ranked by HLADH XP score. The highlighted numbers represent the best scores. The location of fragments compared to NADH is presented in the last column.

Fragments	HLADH		HLADH without Zn <sup>2+</sup>		Similar location to NADH (w/wo Zn <sup>2+</sup> )
	XP Glide	$\Delta G$	XP Glide	$\Delta G$	
NADH	-14.8	-80.6	-15.1	-83.6	-
NAD <sup>+</sup>	-11.5	-92.7	-12.0	-125.7	-
<b>6</b>	<b>-10.9</b>	<b>-62.5</b>	<b>-8.8</b>	<b>-70.9</b>	Yes/yes
<b>9</b>	-10.6	-55.4	-8.6	-69.9	No/no
<b>8</b>	-10.3	-40.5	-8.6	-58.1	No/yes
<b>11</b>	-9.2	-39.8	-6.6	-40.3	No/no
<b>7</b>	-8.7	-53.3	-6.7	-58.9	No/no
<b>12</b>	-8.7	-24.4	-6.1	-22.5	No/no
<b>14</b>	-8.6	-23.6	-7.4	-41.1	No/yes
<b>10</b>	-7.5	-36.1	-7.3	-40.3	No/no
<b>13</b>	-6.5	-41.9	-7.9	-40.4	No/no

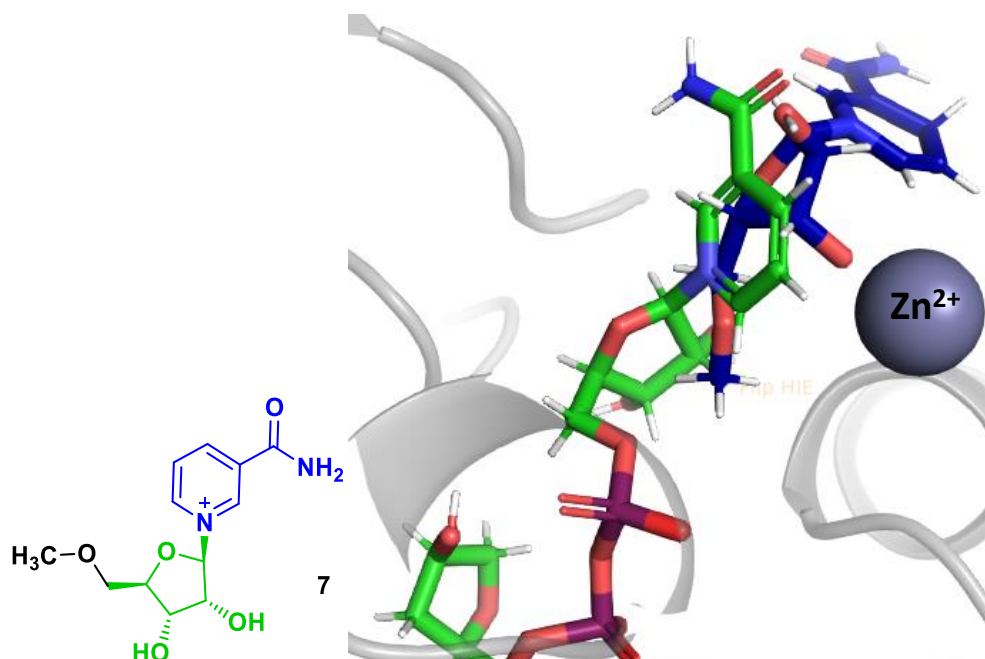
Docking scores of all fragments presented in **Table 3** displayed lower affinity than the entire NAD(H). Fragment **6** which contained a ribose moiety modified with a phosphate group at the primary hydroxyl, presented the lowest free energy binding using both Glide XP and MM-GBSA methods.

The visual analysis showed a general location of the fragment in the end of NAD(P)H pocket facing the active site, which is naturally occupied by the cofactor nicotinamide ring and the pyrophosphates groups. More precisely, nicotinamide fragments (**6 to 8**) were roughly located in the nicotinamide binding region, featuring similar contacts (**Figure 39**).



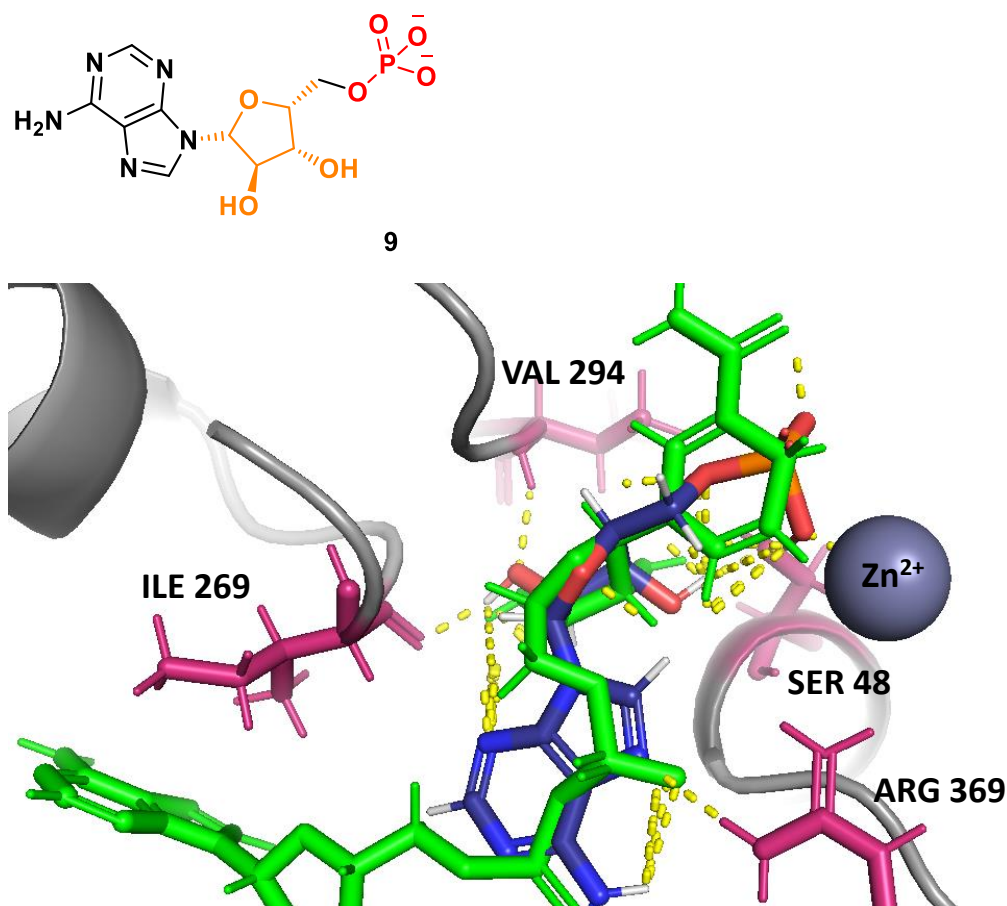
**Figure 39:** Nicotinamide mononucleotide fragment (NMN, fragment 6, blue) superposed to NADH (green). The H-bonds are in dashed yellow lines, while the amino acids (pink) in contact with the ligands are labelled.

When no phosphate groups were present in the ligand structure, a shift of the nicotinamide ring was observed towards a hydrophobic pocket further from the normal position of NADH (e.g. with fragments 7, **Figure 40**). This pocket is positioned on the other side of the  $\text{Zn}^{2+}$ , constituting the binding site of the substrate.<sup>217</sup>



**Figure 40:** The nicotinamide fragment 7 (blue) superposed with NADH (green), slide inside the hydrophobic substrate binding pocket.

On the other hand, in the best scored adenosine fragments (**9** or **11**) the ribose was positioned at the native NADH nicotinamide ribose position, making similar hydrogen bonds with V268, V294 and I269 (**Figure 41**).



**Figure 41:** Example of adenosine monophosphate fragment **9** (blue) superposed to NADH (green), and its interaction residues (pink).

All adenosine fragments were in the ribose and nicotinamide ring area, with few interactions made by their hydroxyl or phosphate group.

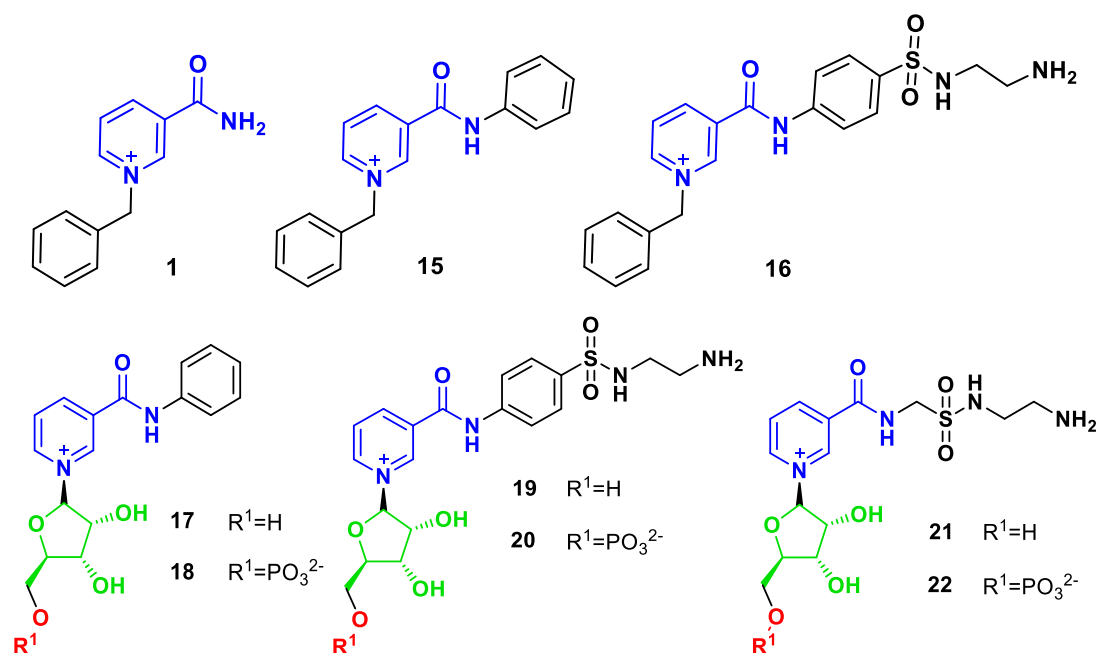
In general, scores obtained with all the fragments were of comparable values. This highlighted a similar affinity, confirmed by the same interactions detected in all the 3D structures. The fragment docking approach gave first insights in the importance of each NAD(P)H part. Adenosine and adenosine 3'-monophosphate showed little affinity for their respective sites. This suggested that adenosine fragments were least responsible for NADH affinity.

Considering the better scores when they were present, the ribose and 5'-phosphate groups seemed to lead to better binding affinity on any of the fragments.

### 3.3.1.5 Docking of NAD mimics from literature

Literature reports suggested that benzyl nicotinamide structures and nicotinamide mononucleotide (NMN, **fragment 6**) were active NADH mimics with HLADH.<sup>104</sup> On the other hand, other reports implied no activity with benzyl nicotinamide in HLADH.<sup>218</sup> The affinity of these mimics for HLADH are still under debate.

Following the results of fragments docking, the nicotinamide part was presented as the important part for binding. Thus, the NADH nicotinamide ring was used as frame for the design of ligands to dock inside HLADH, alongside mimics from Fish and co-workers and M. Basle work (**Figure 42**).<sup>104,109</sup> The oxidised form was applied for all structures, as a slight affinity advantage was shown in the validation of the protocol with the natural cofactors (section 3.3.1.3, page 66).



**Figure 42:** Designed ligands for further docking screening. Structures **1**, **15** and **16** are based on previous work from literature.<sup>104,109</sup> Structures **17** to **22** are based on NMN fragments, supplemented with a sulphonamide bidentate ligand for structures **19** to **22**.

The first set of ligands (**1**; **15**; **16**) were composed of a nicotinamide ring to which a benzyl group was added, based on the previous designs of mimics from Fish and Hollmann work (**Figure 25**, **Figure 26**, page 31).<sup>104,107</sup> These analogues already synthesised and tested for inhibition in our laboratory, gave a reference for the docking results. For ligand **16**, a sulphonamide group was also included, that later should support the transition metal for the imine reduction.<sup>124</sup>

Based on the structure of NMN, a second set of ligands was then created using a ribose and a phosphate group (ligands **17** to **22**, **Figure 42**). The amide group was substituted with a phenyl to highlight the effect of a hydrophobic group in this part of the structure (ligands **17**, **18**). The sulphonamide complex was also added to the structure (ligands **19**, **20** and **21**, **22**). These molecules were designed to test the impact of ribose and phosphates on affinity and the fragments **6** and **8** were kept as reference. All stereochemistry of the ligand structures were kept identical to the co-crystallised NADH cofactor. To investigate the possible impact of docking parameters concerning  $\text{Zn}^{2+}$  interactions on ligands, both HLADH with and without  $\text{Zn}^{2+}$  were used for the docking, applying the docking protocol established earlier.

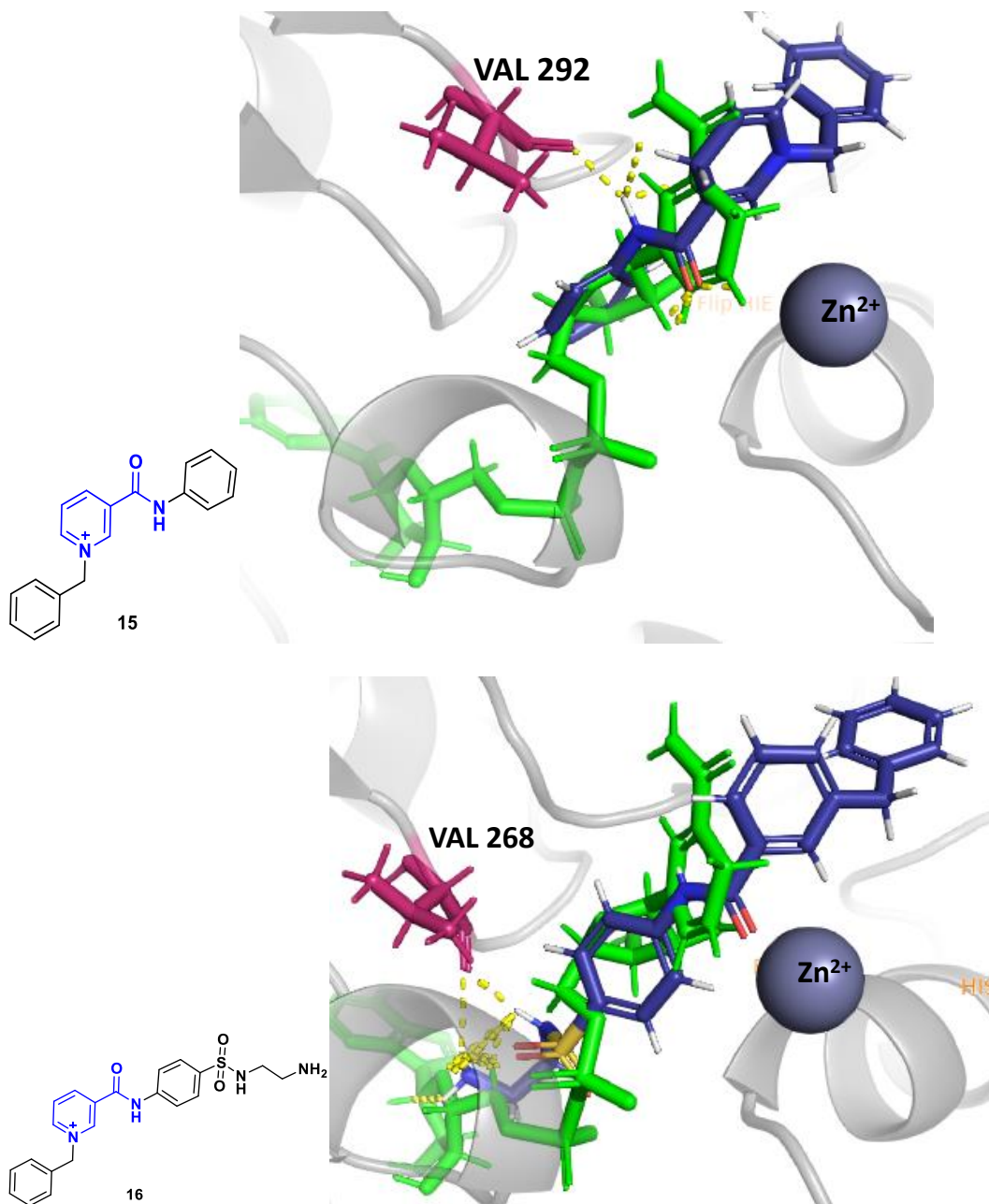
**Table 4:** Presentation of Glide XP ( $\text{kcal mol}^{-1}$ ) and MM-GBSA ( $\Delta\text{G kcal mol}^{-1}$ ) docking scores for the designed compounds and NAD(P)H inside HLADH with and without  $\text{Zn}^{2+}$ . The ligands are ranked by XP results in HLADH. The highlights represent the best scores. The location of ligand nicotinamide ring is compared to NADH for both docking inside HLADH and HLADH without  $\text{Zn}^{2+}$ .

Ligands	HLADH		HLADH without $\text{Zn}^{2+}$		Similar position to NMN (w/wo $\text{Zn}^{2+}$ )
	XP Glide	$\Delta\text{G}$	XP Glide	$\Delta\text{G}$	
NADH	-14.8	-80.6	-15.1	-83.6	-
NAD <sup>+</sup>	-11.5	-92.7	-12.0	-125.7	-
<b>18</b>	<b>-9.9</b>	-53.5	<b>-10.5</b>	-65.7	yes/yes
<b>19</b>	-9.8	-65.2	-7.7	-81.9	no/no
<b>17</b>	-9.7	-63.1	-10.2	<b>-82.3</b>	no/yes
<b>20</b>	-9.2	-46.2	<b>-12.2</b>	<b>-81.8</b>	no/yes
<b>22</b>	-8.4	-42.7	-9.2	-70.9	yes/yes
<b>6</b>	-6.8	<b>-67.3</b>	-8.6	-70.2	yes/yes
<b>15</b>	-6.7	-43.1	-5.9	-70.4	no/yes
<b>16</b>	-6.2	-50.3	-2.8	-77.9	no/no
<b>8</b>	-4.1	-62.5	-6.5	-71.4	yes/yes
<b>21</b>	-4.0	-57.0	-5.8	-41.0	yes/no
<b>1</b>	-3.7	-47.4	-4.2	-62.8	yes/yes

The docking results showed similar Glide XP ranking between both HLADHs while MM-GBSA scores resulted in higher values for HLADH without  $\text{Zn}^{2+}$ .

The benzyl-modified ligands (**1**; **16**) resulted in significantly lower binding affinity inside both HLADH crystal structures compared to all other molecules, as shown by

their lowest rank with Glide XP (**Table 4**). The location of the nicotinamide ring was identical to NADH ring; however, the benzyl group was placed inside the substrate pocket. In addition, the ligands displayed a limited number of specific bonds (e.g. H-bonds) with the protein, mainly between the sulphonamide and the amide group with valine residues in the pyrophosphate binding area (**Figure 43**). However, this situation was not acceptable for the final objective to construct a binding metal catalyst, due to the sulphonamide support taking the space of the anchoring part and thus positioning the metal toward the exit of the cofactor binding pocket.

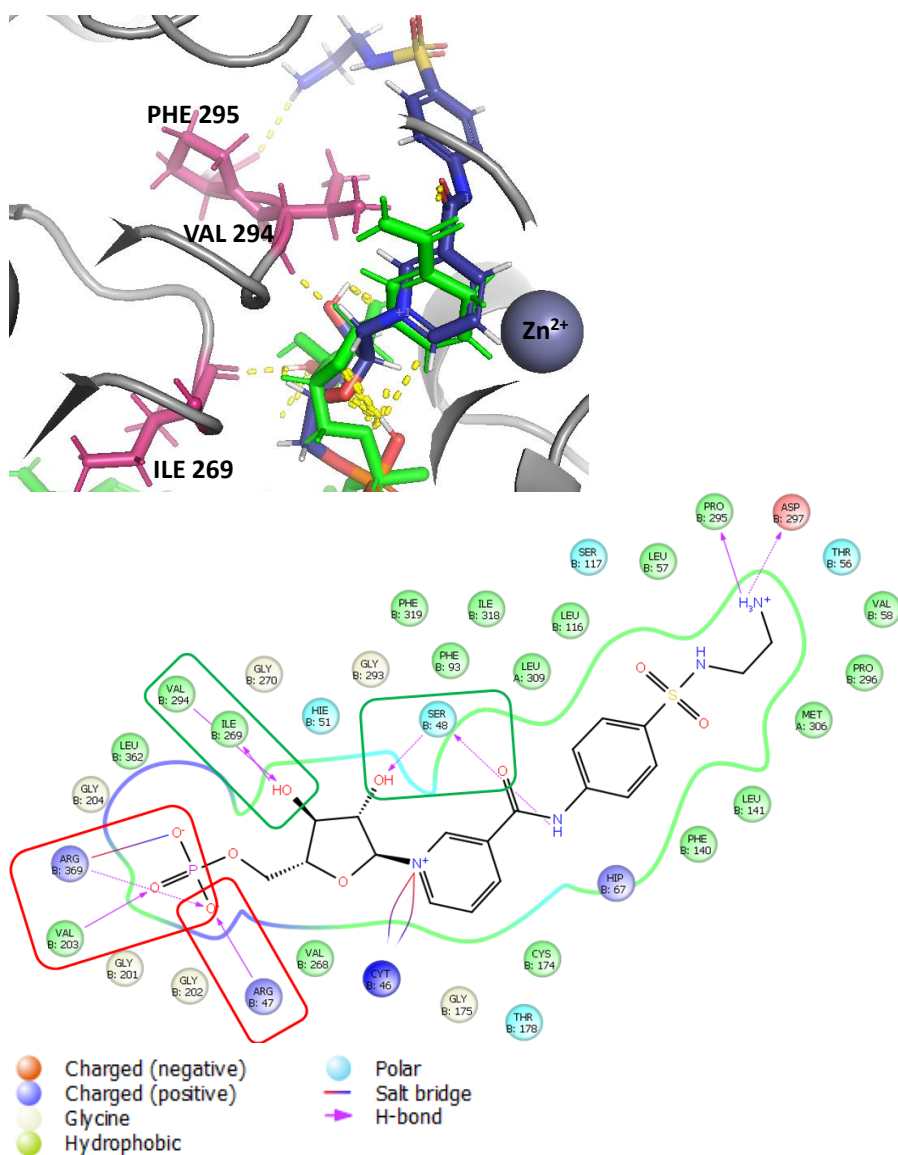


**Figure 43:** Best scored docking poses of ligand **15** and **16** (blue) superposed to NADH (green). The benzyl was inserted in the substrate hydrophobic pocket. The ligands bound to valines (pink) with their amide and sulfonamide group.



The lack of contacts and a wrong position of the ligand prevented a strong affinity toward HLADH and a right positioning to a catalytic activity. These docking results confirmed previous conclusions of inhibition assays from M. Basle,<sup>109</sup> along with the work of Hollmann and co-workers<sup>218</sup> who suggested a low affinity for the synthesised ligands toward HLADH.

The second set of ligands (**17** to **22**) was docked to establish the ribose and phosphate groups importance in the binding location and affinity. The structures were ranked at the top with the best XP and MM-GBSA scores (**Table 4**). When looking at the poses, all the ligands inserted their sulphonamide in the substrate hydrophobic pocket ahead of NADH pocket, using the phosphate and the ribose for H-bonds interactions (**Figure 44**).



**Figure 44:** 3D structure and corresponding interaction diagram of ligand **20** (blue). Phosphate and ribose made interactions that reversed the structure's orientation, inserting the sulphonamide group



inside the substrate pocket, ahead of the NADH binding pocket. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, Ile: isoleucine, Leu: leucine, Lys: lysine, Met: methionine, Phe: phenylalanine, Pro: proline, Ser: serine, Thr: threonine, Val: valine.

The docking poses of the best structures suggested again a phosphate and sugar impact in affinity, confirming the previous NADH fragments docking results. The presence of a ribose and a phosphate completely reversed the ligand structure orientation by positioning the sulphonamide inside the substrate hydrophobic pocket, contrary to ligand **1** or **15** whose benzyl anchoring part was in the substrate pocket.

To conclude on these first screenings inside HLADH, all best fitting ligands showed similar features:

- Ribose and phosphate groups are important for the binding.
- Several functional groups (carboxyl, amine, and hydroxyl) in the structure's core make hydrogen bonds all along the binding pocket (especially in ribose/pyrophosphate area with valine and arginine residues).
- A hydrophobic part (aromatic rings, aliphatic chain, halogens) docks in the substrate pocket, lined with hydrophobic residues.

These findings led to a series of new questions arising: can we use the substrate pocket for ligand binding, or is this area too important for the catalytic reaction? When the substrate is present, is there enough space remaining to include a metal? New docking experiments should be conducted with the presence of a substrate preventing any docking inside its pocket, to confirm the previous interaction hypotheses and optimize them.

### 3.3.2 *Docking with HLADH co-crystallised with substrate*

A crystallized structure of HLADH complexed with NAD<sup>+</sup> and 2, 3, 4, 5, 6-pentafluorobenzyl alcohol (PDB 4DWV) was prepared with and without Zn<sup>2+</sup> ion.

By avoiding occupancy of the substrate binding pocket, the study of this crystal structure might provide more information on major contacts inside the cofactor binding site.

### 3.3.2.1 Docking screens for NAD(P)H cofactors and NADH fragments

The docking protocol was verified in the previous section (3.3.1.1, page 63) for this new HLADH structure, and the RMSD scores validated the protocol. Similar steps to the previous docking were then applied with the docking of NAD(P)H cofactors and their fragments.

**Table 5:** Glide XP (kcal mol<sup>-1</sup>) and MM-GBSA ( $\Delta G$  in kcal mol<sup>-1</sup>) results inside HLADH with substrate, ranked by XP score. The highlighted numbers represent the best scores. The location of fragments is compared to the co-crystallised NAD<sup>+</sup> in the last column. HLADHsubst refers to the crystal structure containing a substrate (PDB 4DWV).

Fragments	HLADHsubst		HLADHsubst without Zn <sup>2+</sup>		Similar location to NAD <sup>+</sup> (w/wo Zn <sup>2+</sup> )
	XP Glide	$\Delta G$	XP Glide	$\Delta G$	
<b>NADH</b>	-14.6	-60.8	-14.1	-80.5	-
<b>NADPH</b>	-12.9	-95.2	-14.6	-90.2	-
<b>NAD<sup>+</sup></b>	-11.4	-101.2	-13.1	-119.9	-
<b>NADP<sup>+</sup></b>	-11.9	-77.9	-14.5	-73.4	-
<b>6</b>	<b>-8.9</b>	<b>-50.6</b>	<b>-10.2</b>	-94.2	yes/yes
<b>9</b>	-8.2	-24.3	-4.8	-33.8	yes/yes
<b>8</b>	-7.8	-47.9	-10.1	-63.9	yes/yes
<b>7</b>	-7.8	-34.9	-6.4	-93.9	no/yes
<b>10</b>	-6.8	-38.9	-7.2	-42.3	no/no
<b>11</b>	-6.7	-28.1	-5.6	-45.5	no/no
<b>13</b>	-7.1	-35.5	-5.4	-36.0	no/no
<b>12</b>	-6.4	-38.3	-7.4	-41.8	no/no
<b>14</b>	-6.9	-46.8	-7.5	-34.3	yes/no

This new docking showed similar conclusions as the earlier docking in HLADH without the substrate, in terms of score and visual analysis:

- Lack of discrimination between NADH and NADPH, and between their reduced and oxidised forms.
- NADH fragments location was restricted to middle and top areas of the binding pocket and fragment **6** (NMN) was clearly selected as the best score (**Figure 38**).
- Similar to the NAD<sup>+</sup> co-crystal, contacts were made by the ribose and phosphate groups with arginine and valine residues, confirming the importance of the nicotinamide area in the cofactor pocket.

### 3.3.2.2 Docking screens for designed ligand

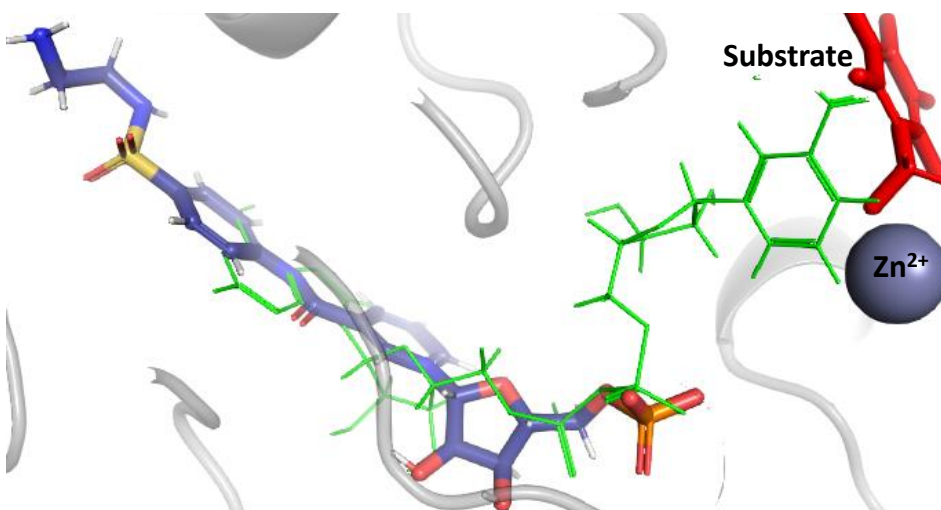
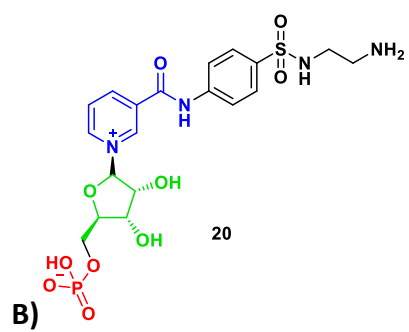
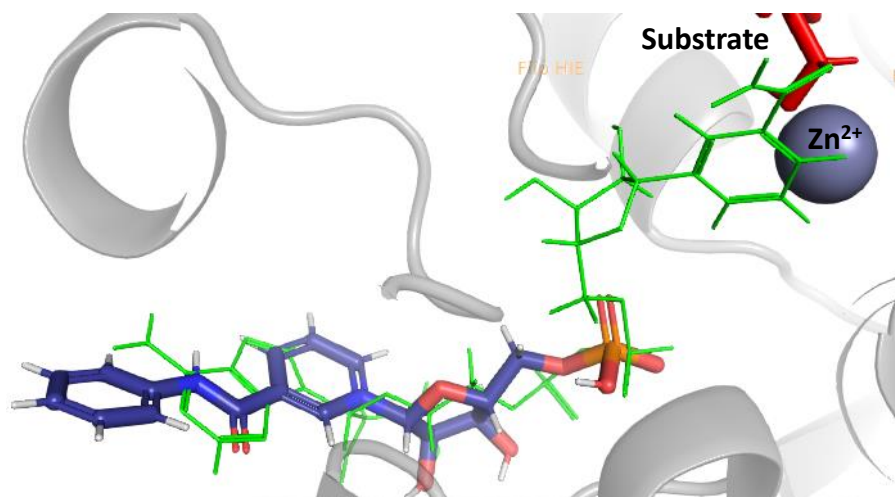
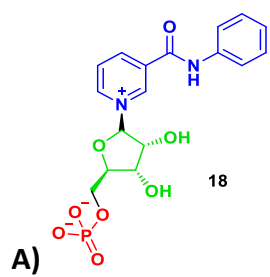
The same ligands (**1** to **22**; **Figure 42**) were docked again inside the binding site in the presence of the co-crystallised substrate (see **Table 6**).

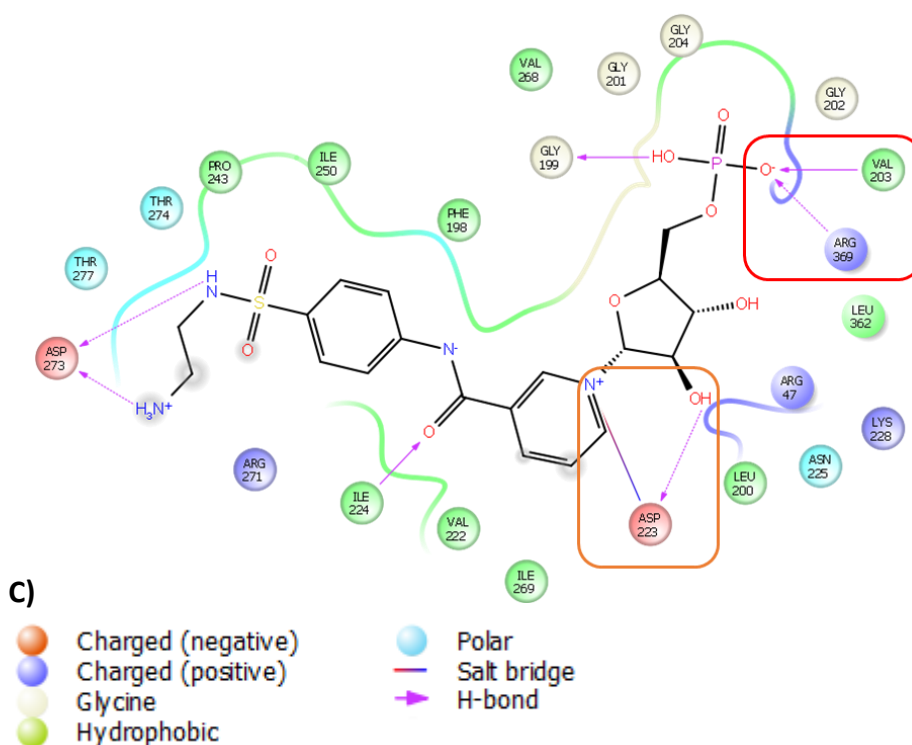
**Table 6:** XP (kcal mol<sup>-1</sup>) and MM-GBSA ( $\Delta G$  in kcal mol<sup>-1</sup>) docking results, for designed ligand inside HLADH including co-crystallised substrate with and without Zn<sup>2+</sup>. The ligands are ranked by XP results and the highlights represent the best scores. The location of ligand nicotinamide ring is compared to the NAD<sup>+</sup> co-crystal. HLADHsubst refers to the crystal structure containing a substrate (PDB 4DWV).

Ligands	HLADHsubst		HLADHsubst without Zn <sup>2+</sup>		Similar orientation than NMN
	XP Glide	$\Delta G$	XP Glide	$\Delta G$	
NADH	-9.8	-60.8	-14.3	-54.1	-
NAD <sup>+</sup>	-11.4	-101.2	-13.1	-119.9	-
<b>22</b>	-9.0	-58.0	-9.0	-35.3	No/no
<b>20</b>	-8.2	-36.8	-8.0	-29.2	No/no
<b>18</b>	-8.2	-48.8	-10.4	-93.7	no/no
<b>17</b>	-7.7	-48.8	-8.1	-93.5	Yes/yes
<b>21</b>	-7.3	-36.1	-6.7	-55.2	No/no
<b>16</b>	-7.1	-44.8	-6.6	-62.2	No/no
<b>15</b>	-5.0	-36.2	-1.8	-58.6	No/no
<b>1</b>	-4.5	-41.1	-5.2	-77.3	Yes/yes
<b>19</b>	-4.3	-43.5	-6.9	-48.6	No/no

The ligand structures with phosphates and ribose groups scored again best with both Glide XP and MM-GBSA (**17**, **18**, **22**), while the benzyl-substituted nicotinamide structures **1** ranked among the last. This again confirmed previous docking results with HLADH alone.

When a phenyl or a sulphonamide group was added to the nicotinamide ring position (**18**, **20**), compounds reversed their orientation towards the adenine part location while keeping the ribose and phosphate contacts (**Figure 45**). This was probably due to the additional phenyl ring not able to accommodate inside the substrate pocket. Therefore, the residue D223 in the adenine section of the cofactor pocket interacted with the nicotinamide pyridinium by making a salt bridge.





**Figure 45:** **A)** Docking poses of ligands **18** and **20** (blue), compared to crystal structure **NAD<sup>+</sup> 2** (green), in presence of the substrate (red). **B)** Pose of ligand **20** (blue) superposed with **NAD<sup>+</sup> 2** (green). **C)** Interaction diagram represents ligand **20** contacts with the amino acids inside NADH binding pocket. The red box represents interaction similar to NADH pyrophosphate and the orange box the NADH adenine ribose interactions. Arg: arginine, Asp: aspartic acid, Asn: asparagine, Gly: glycine, Ile: isoleucine, Leu: leucine, Lys: lysine, Phe: phenylalanine, Pro: proline, Thr: threonine, Val: valine.

No significant differences appeared for structures without hydrophobic substituents inside HLADH with substrate, where they remained in a NADH like position. However, an inversion was observed for bigger and more hydrophobic structures (ligands **16** to **18**) which arranged their hydrophobic part toward the outside of the pocket, while keeping the phosphate and ribose interactions.

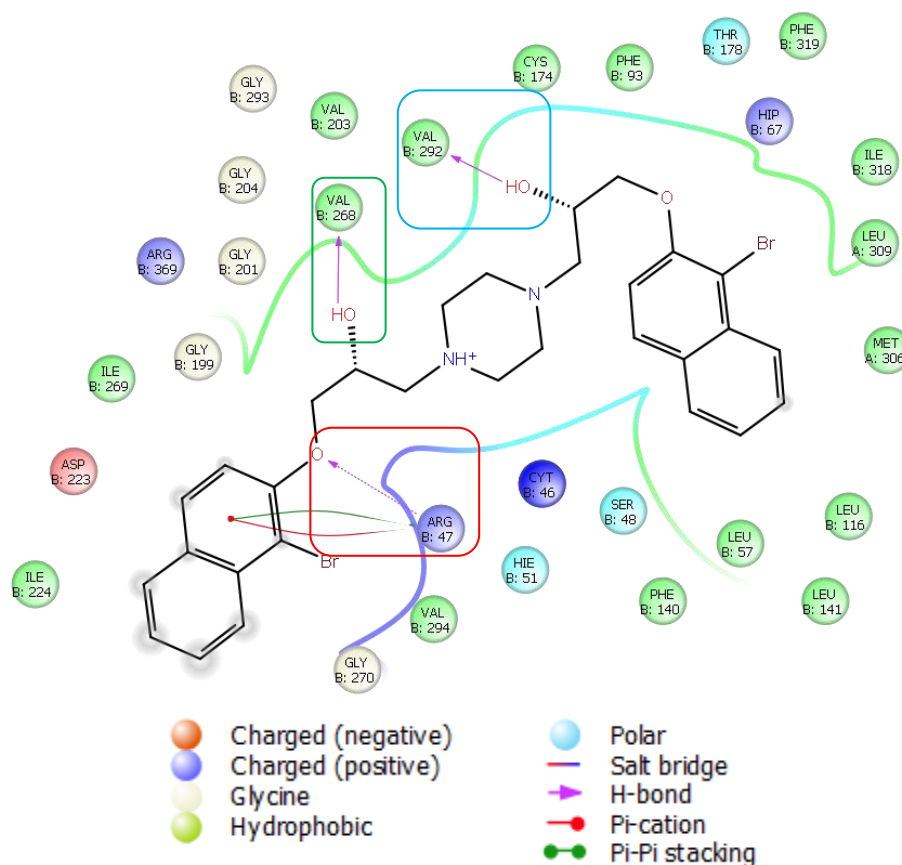
This docking with substrate confirmed the previous conclusion settled: the ribose and the phosphate seem to be the decisive groups in NADH structure, while several links with arginine (R369 and R47) and valine residues (e.g. V294, V292 and V268) are of importance for a good affinity. Moreover, supplementary hydrophobic substituents linked to the nicotinamide ring must be avoided for their risk of insertion into the substrate binding pocket. This might block the substrate normal binding and change the right orientation of the ligand toward the substrate. Finally, the nicotinamide ring showed no interaction in all the different docking experiments. This heterocycle might thus not be essential for the binding inside HLADH.

### 3.3.3 Libraries docking

A virtual screening study was conducted using libraries of ligands to enhance the variety of structures docked inside HLADH. This will highlight the structural characteristics of a ligand with affinity for the enzyme cofactor pocket, and the residues used for the interactions. Around 8000 structures of available chemical compounds and their conformers were docked within both HLADH and HLADH with substrate. The online available Schrodinger library of lead like ligands and the ZINC library were used (ZINC12; Glide Drug-like Ligand Decoy sets).<sup>219</sup>

#### 3.3.3.1 Library screening in HLADH

The first ten best binding ligands showed similar XP and MM-GBSA values as the NADH cofactor inside HLADH. They docked with a MM-GBSA score between -89.0 kcal mol<sup>-1</sup> and -59.4 kcal mol<sup>-1</sup>, where NADH was of -80.6 kcal mol<sup>-1</sup> and NAD<sup>+</sup> -92.7 kcal mol<sup>-1</sup>.



**Figure 46:** Interaction diagram of best MM-GBSA scoring ligand in HLADH cofactor binding pocket. Similar interactions as phosphate (red box), ribose (green) and nicotinamide (blue) groups of NADH were spotted. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Met: methionine, Phe: phenylalanine, Ser: serine, Thr: threonine, Val: valine.

In a general visual analysis, the docking of libraries inside HLADH again validated the previous conclusions. The best ranked ligands shared two important features in their structures: a hydrophobic part with naphthalene or aliphatic chain and a hydrophilic centre composed of amino, carboxyl or hydroxyl groups.

The hydrophobic elements were always positioned in the hydrophobic substrate pocket (**Figure 46**) while the hydrophilic groups formed several interactions with residues in the cofactor pocket. A recurrence of three or four links was spotted between carboxyl/hydroxyl groups and R369, R47 and/or V203, V292.

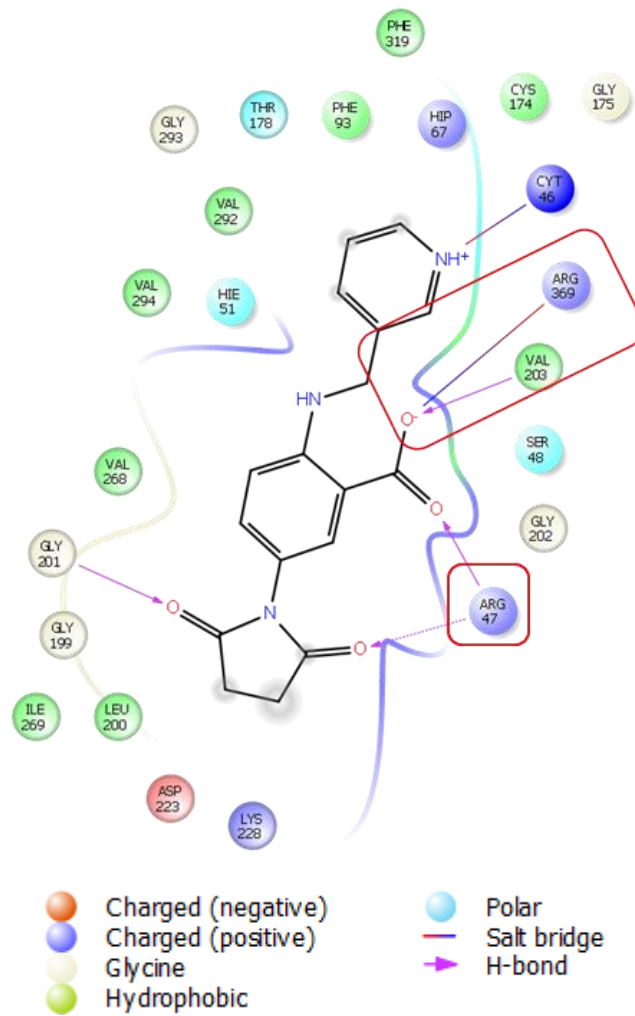
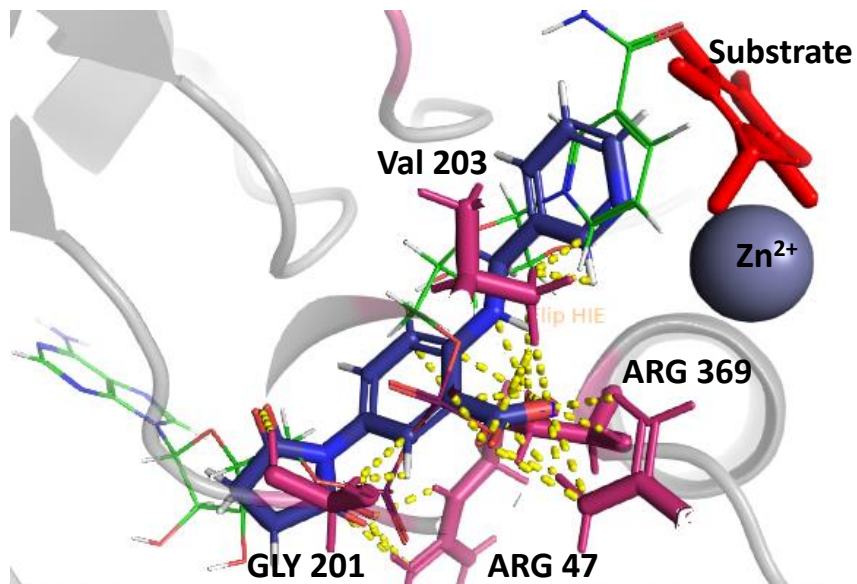
On the other hand, the lowest ranked structures were mainly composed of hydrophobic groups with aromatic rings and heterocycles.

#### 3.3.3.2 Large screening in HLADH with substrate

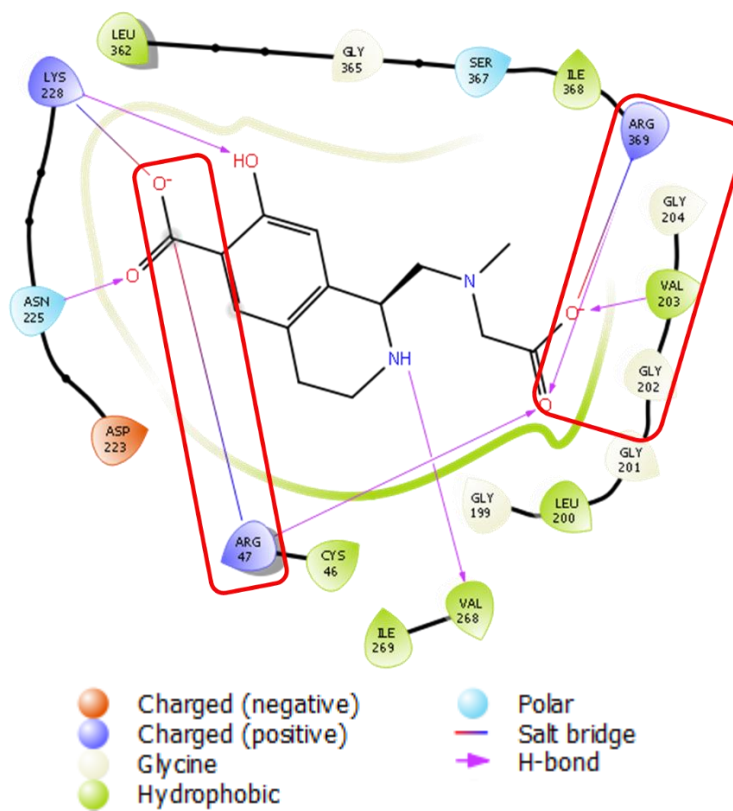
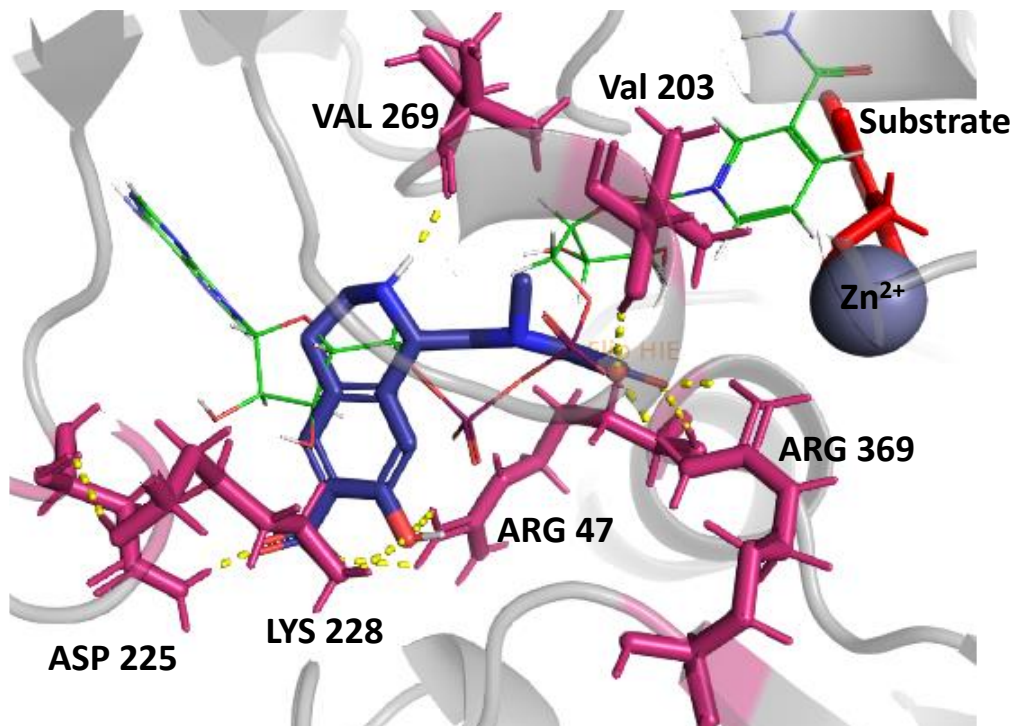
The same ligands libraries were screened inside HLADH with substrate. The docking showed a good correlation between the two HLADH structures with and without  $\text{Zn}^{2+}$ ; the score values and the structures were found similar (hydrophilic groups and heterocycles).

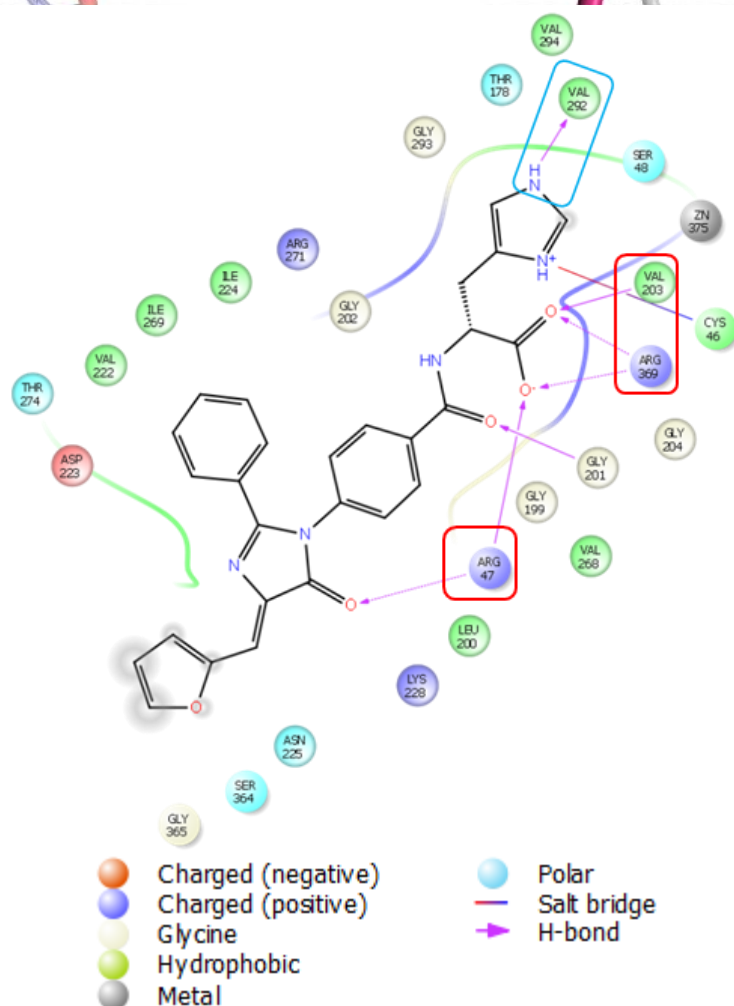
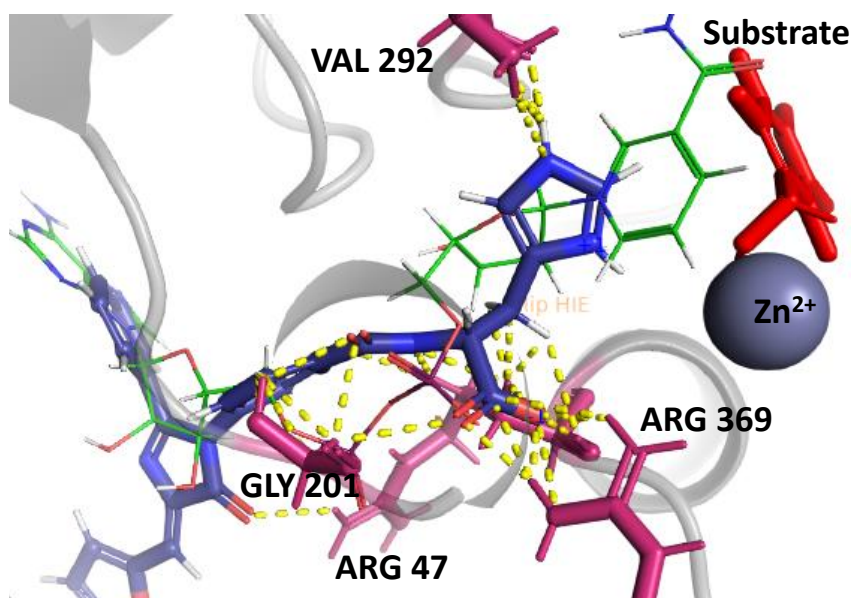
Compared to HLADH without substrate, lower MM-GBSA scores were obtained (between  $-38.4 \text{ kcal mol}^{-1}$  and  $-67.5 \text{ kcal mol}^{-1}$ ).

The first visual impact of these ranking differences was seen in the ligand's structure shape, where hydrophilic groups were even more frequently present. For example, hydrophobic groups frequent in previous HLADH alone docking were generally replaced by heterocycles (**Figure 47**). The substrate pocket being occupied, the hydrophobic part was reduced in the selected structures. The ligands docked inside the cofactor pocket only, by contacts from the nicotinamide and phosphate area (**Figure 47**). Almost no interactions were present in the adenine area.





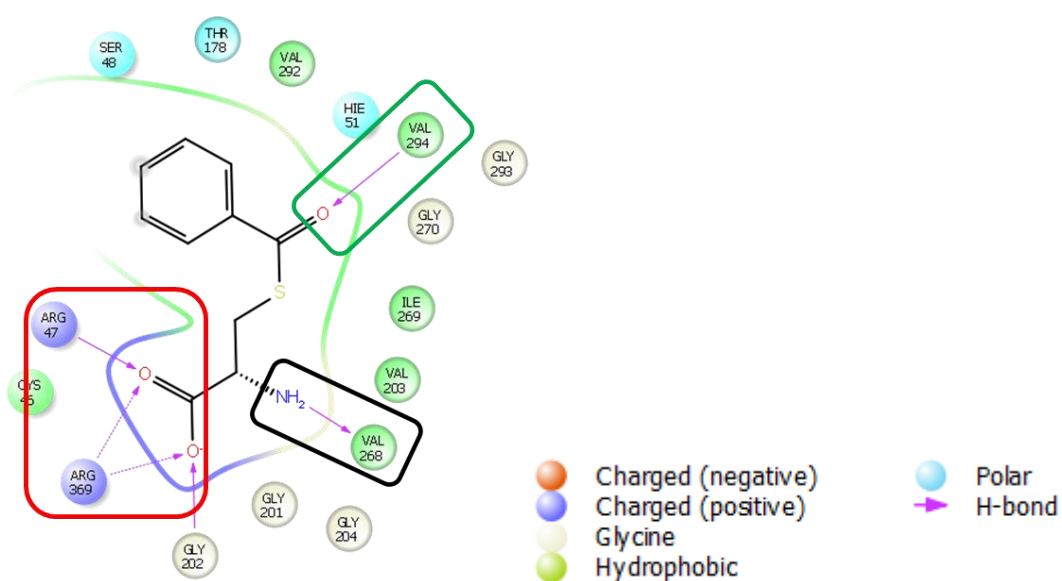
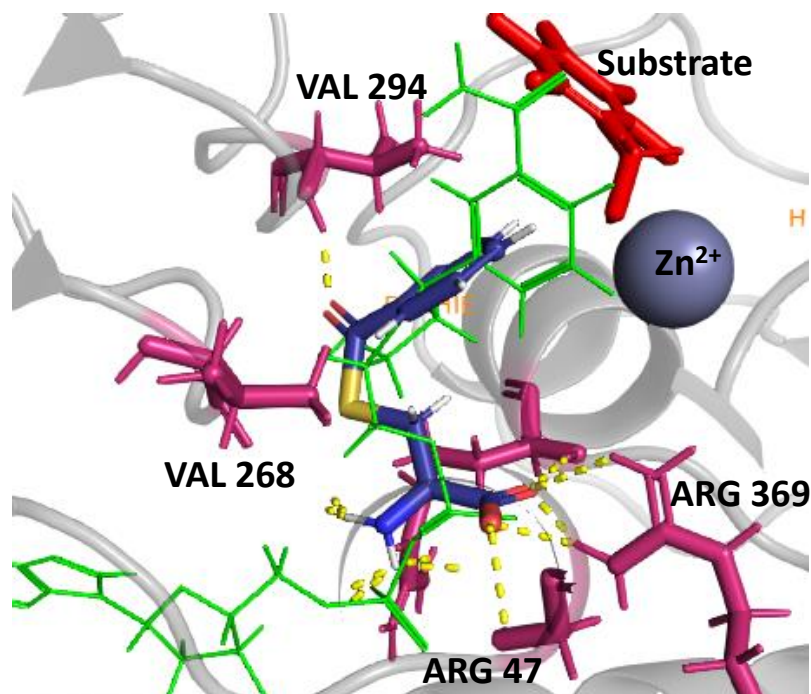




**Figure 47:** Selection of the ligands poses from libraries with the best XP and MM-GBSA scores, and their corresponding interaction diagrams inside the cofactor binding pocket of HLADH. Ligands (blue) are superposed to  $\text{NAD}^+$  (green) with substrate (red). In the diagrams, the similar interactions as the NADH pyrophosphate group are represented in red boxes and the nicotinamide in blue. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Phe: phenylalanine, Ser: serine, Thr: threonine, Val: valine.

Schrodinger and ZINC libraries also provide fragments sets of ligands, which can be used for a more precise screening of important areas.

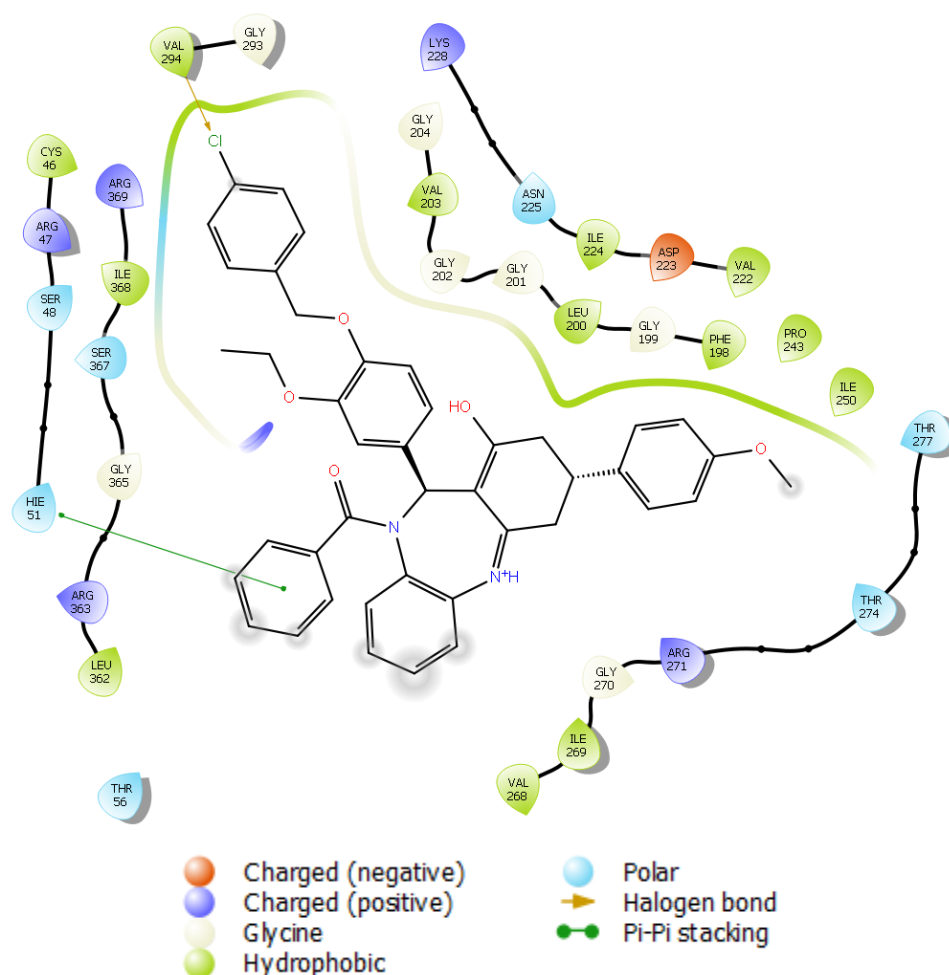
The analysis of fragments docking confirmed the previous findings: the general localisation of the fragments was identical to the pyrophosphate position of the co-crystallised NADH (**Figure 48**). They also formed similar interaction with V294, V203 and R47, R369.



**Figure 48:** Pose and interaction diagram of the best scoring fragment from the online libraries (blue). The interactions mimicked NAD<sup>+</sup> ribose (green box) and phosphate interactions (red box). Arg: arginine, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Ser: serine, Thr: threonine, Val: valine.

In HLADH with substrate, all best poses always made interactions with V203 and V292, V268 and V294 in addition of two hydrogen bonds with R47 and R369 inside the cofactor pocket. Other less frequent links could also be spotted with glycine, cysteine, or histidine residues. Additional extra links with residues were observed compared to the docking in HLADH alone, where the hydrophobic cleft of the substrate added some affinity for the ligand docking.

The lowest ranked ligands obtained XP scores around 0.796 and 1.134 (**Figure 49**) and were typically hydrophobic structures, with few interactions.



**Figure 49:** An example of the lowest ranked structure from Schrodinger library, with few interactions. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Phe: phenylalanine, Pro: proline, Ser: serine, Thr: threonine, Val: valine.

### 3.3.4 Conclusion of the first docking study in HLADH

Despite the variety of ligands tested in this first docking studies, similar structures and interactions were found to be an advantage for the affinity inside HLADH. Those

docking studies provided insights into the NADH structural motifs and the cofactor pocket residues relevant for an efficient binding:

- The adenosine moiety and the nicotinamide ring did not improve binding significantly.
- The phosphate group and the ribose nicotinamide seemed to be the most important groups in the NADH structure for high affinity, by building strong H-bonds.
- Key residues in the cofactor pocket for strong affinity include mainly arginines (R47, R369) and valines (V203 and V294). Two interactions frequently observed with the arginine residues implied even better scores and ranking for the ligands.<sup>213</sup>
- The substrate pocket was useful for hydrophobic contacts. It may bring higher affinity and stability to the complex catalyst-enzyme. For example, a hydrophobic area in the ligand could relocate its metal closer to the substrate. On the other hand, this means such ligands would occupy the natural substrate binding site, which might prevent further substrate binding necessary for the anticipated catalysis.

### 3.3.5 *Improved docking strategy and design of small anchor library*

#### 3.3.5.1 Development of a new computational screening method

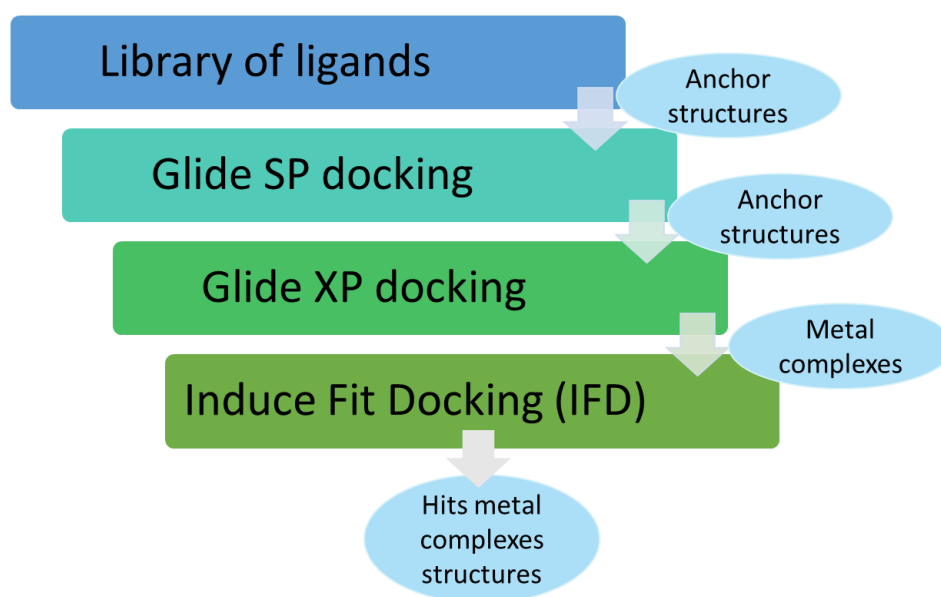
Based on the structural characteristics observed during the first docking experiment, a small anchor library was designed. To screen this new library and define a smaller set of lead structures, a second and improved docking strategy was developed.

In the pharmaceutical industry, computational docking is generally used to identify inhibitors of proteins by favouring best binding structures. Hence, the affinity between a protein and its ligand is the principal parameter studied.<sup>162</sup> In that way, during the first part of this computational study, the structures were ranked based on their binding energies.

However, for this project objective, not only binding affinity but exact binding orientation were of major importance to find a suitable hit anchor for a metal complex catalyst. The anchor must place the metal exactly towards the substrate

pocket to deliver a catalytic activity. During these new docking studies, precision in the affinity, the location and the orientation of the structures must be considered.

The new scheme of computational docking calculation started on structures without the transition metal, by the traditional Glide SP and XP docking screenings (**Figure 50**). These experiments delivered a raw filtering of the small library of anchor structures. The change in the protocol was the final step with the introduction of Induced Fit Docking (IFD) treating the whole structures of ligands (anchor and metal complex). Indeed, Glide XP or MM-GBSA used in the previous docking study were not able to handle the Cp\* metal complex inside the enzyme. A higher flexibility of the protein scaffold was needed in order to allow the Cp\* group insertion at the interface between cofactor and substrate binding sites. Therefore, MM-GBSA docking was replaced by the more accurate IFD method.<sup>207</sup>



**Figure 50:** The protocol designed for the second set of docking experiments, with a first rough screening using Glide score and a refinement using IFD calculation with some enzyme flexibility in the final step.

In addition to the Glide XP score representing the affinity of ligands for HLADH, a new scoring was to be defined, characterising the orientation and location of ligands. Filters were created to evaluate various poses and explain the choice of structures among the ligand libraries. The location and the orientation of the ligand inside the HLADH binding site were evaluated as follows:

Location: **(x3)**

- 0: not in the pocket

- 1: in the entrance (adenine part)
- 2: in the adenine-phosphate area
- 3: in the pyrophosphate area
- 4: in the nicotinamide ribose area
- 5: perfectly follow the NAD area

#### Orientation compared to NADH: **(x4)**

- 0: turned upside down
- 2: when several poses of ligand with some in reverse and some in similar position to NAD
- 5: same orientation as NAD

#### Energy binding scores (XP Score) **(x2)**

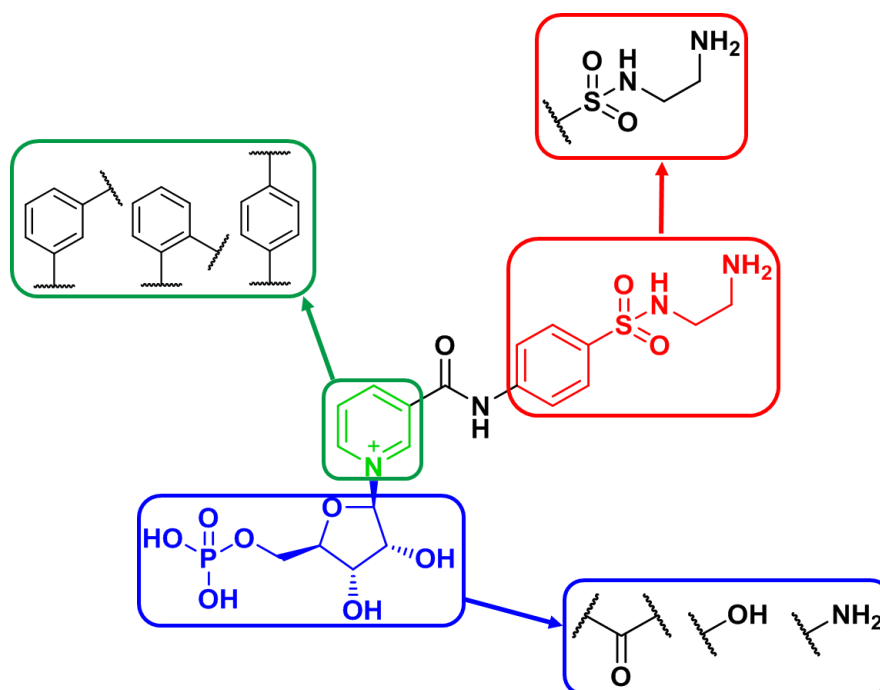
The three factors were also ranked for their importance and impact. Thus, the pose of the structure was the most important parameter and both location and orientation were privileged over the affinity approximation. The resulting final score for each structure allowed a ranking and selection of several potential hit metal catalysts.

HLADH with substrate crystal structure and lacking  $\text{Zn}^{2+}$  was used (PDB 4DWV).

#### 3.3.5.2 Design of a ligand library

In the first set of docking calculations, phosphate and ribose groups were found the most important part of NADH for maximising binding affinity, as they form several H-bonds with arginine and valine residues. On the other hand, the adenine part and the nicotinamide ring were considered not essential for the binding affinity.

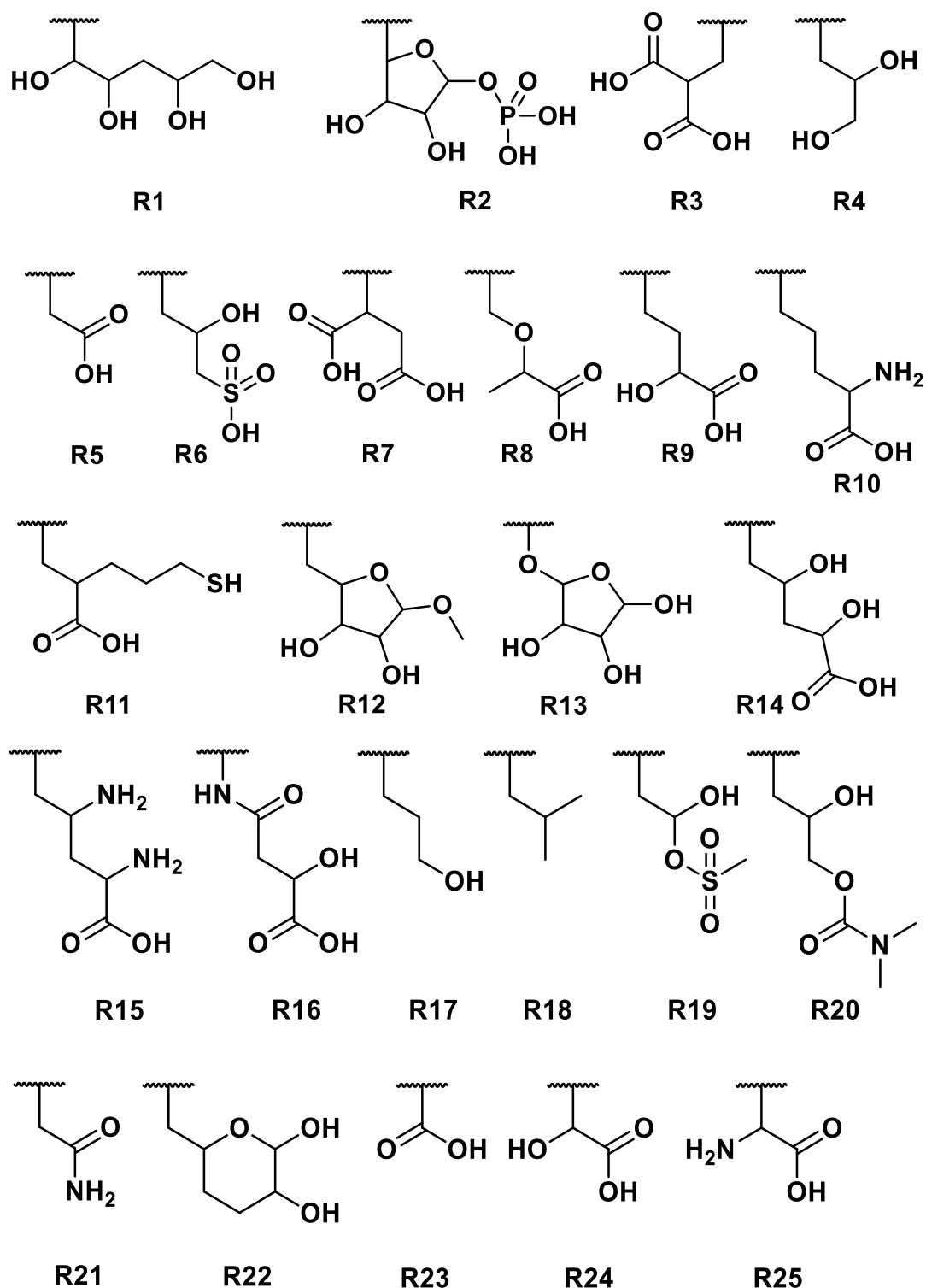
Based on these results, functional groups that mimic or strengthen the interactions between phosphate/ribose group and residues like R47, R369, V203 and V268 were selected. **(Figure 51).**



**Figure 51:** The prototype of the new high affinity ligand structure was based on the nicotinamide part of NAD(P)<sup>+</sup> cofactor. The phosphate and ribose represent the anchoring part and can be replaced by shorter and simpler hydrophilic groups (blue). A sulphonamide bidentate metal ligand with and without its typical aromatic ring (red), is inserted at various positions on the nicotinamide ring.

In view of the first docking results, a new anchor structure library was created (**Figure 52**). Hydroxyl and carboxylic acid groups were used as principal groups to mimic the phosphate and ribose part of NADH. The strategy followed for the anchoring part design begun by a polyol structure, which mimics the two hydroxyl groups of the nicotinamide ribose. Modifications were then brought to the length, position, number, and nature of the functional groups. For example, carboxylic acids (as mimics of the phosphate) were added in various positions and in combination with other amino or hydroxyl groups. Finally, the original nicotinamide mononucleotide fragment was also added and slightly modified (ligand **R2**, **R12**, **R13**). The stereochemistry was kept unrestricted.

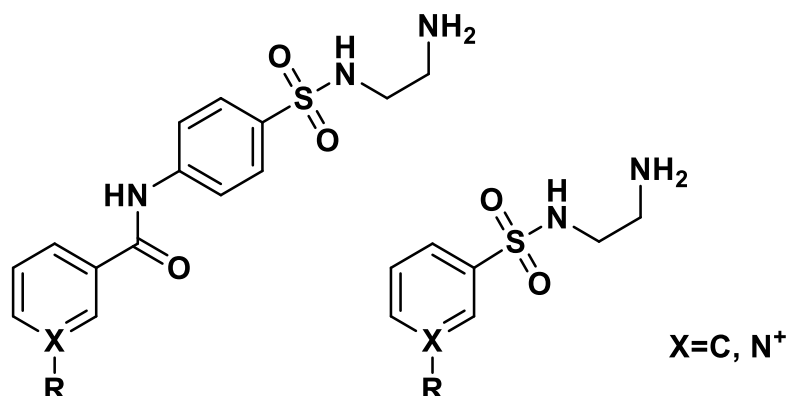




**Figure 52:** Selection of the functional groups (R) composing the new library of ligands. They were created using variation of structure length and hydrophilic groups combinations.

The literature models for efficient transition metal complexes in the catalysis of imine reduction included one aromatic ring.<sup>124</sup> Therefore, these various new functional groups were inserted into structures composed of two rings (pyridine and aromatic ring) as well as on a single ring, to evaluate the impact of structures length on affinity

(**Figure 53**). As highlighted earlier, the nicotinamide demonstrated little effects on the affinity. Thus, the nature of the ring was varied between a pyridine and a phenyl. Finally, the structures were designed with all possible insertion types on the rings (ortho, meta and para position).

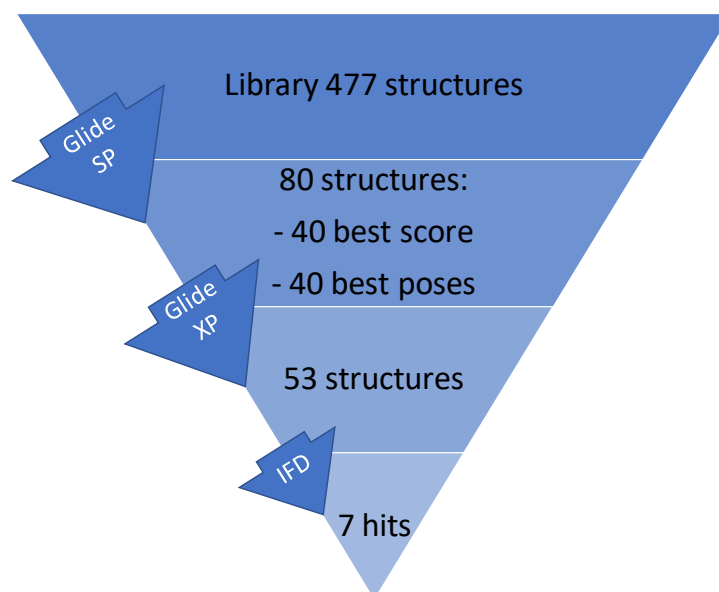


**Figure 53:** Model of ligand structures for the second docking screening, where **R** represents the various substituents (see **Figure 52**) and both phenyl and pyridine rings were used. The sulphonamide served as support for the metal complex and was inserted in ortho, meta or para position in the aromatic ring.

These modifications led to a library composed of 477 newly designed ligands.

#### 3.3.5.3 Glide SP and XP library docking

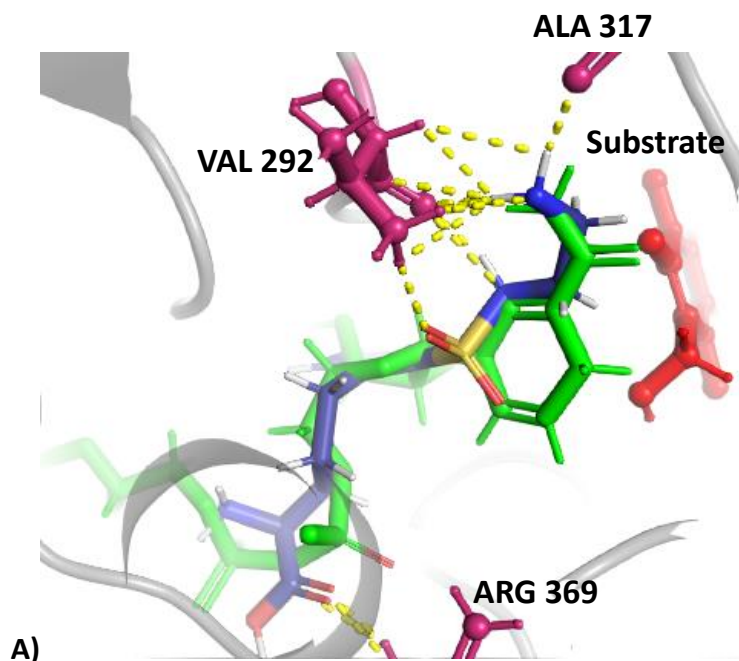
Following the anchor library creation, the docking experiment started with a raw filtering with Glide SP and Glide XP methods. The structures docked at these stages were composed of only the anchor and the sulphonamide metal support (**Figure 53**). The aim of this docking approach was to reduce the number of potential ligands used as anchors, guided by their affinity scores and poses. A suitable pose was defined by a similar orientation and location as NADH inside the pocket. This resulted in a first overview of the **R** substituents impact on the affinity for HLADH.

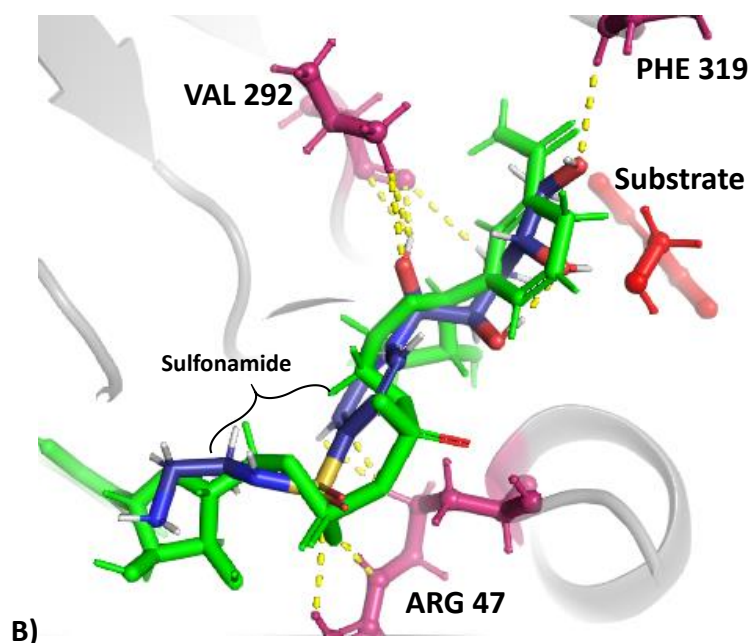


**Figure 54:** General process of the Glide SP / Glide XP / IFD created protocol and the selection of ligands structures at each step.

After the Glide SP filtering, the 40 best scoring ligands were selected, then the 40 best ligand poses were retained and were tested in the Glide XP docking (**Figure 54**).

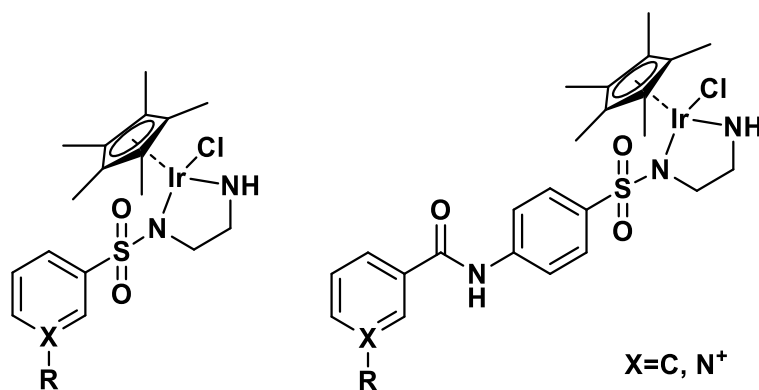
**Figure 55** represents an example of an acceptable and a rejected pose.





**Figure 55:** Superposition of **A)** Anchor **R10** (blue) with XP score of  $-10.3 \text{ kcal mol}^{-1}$  in an efficient pose, with NADH (green) and **B)** Anchor **R1** (blue) with XP score of  $-10.7 \text{ kcal mol}^{-1}$  in a reversed pose, with NADH (green). The sulphonamide would locate the metal toward the solvent.

From the Glide XP resulting ligands, the scores and the poses were manually checked for all the 80 structures. Only the ligands showing an acceptable orientation and location, suitable for the catalytic activity, were selected. Therefore, from the 80 ligands ran through Glide XP docking, 53 ligands were selected (**Figure 54**) and the metal complex IrCp\* was added (**Figure 56**).



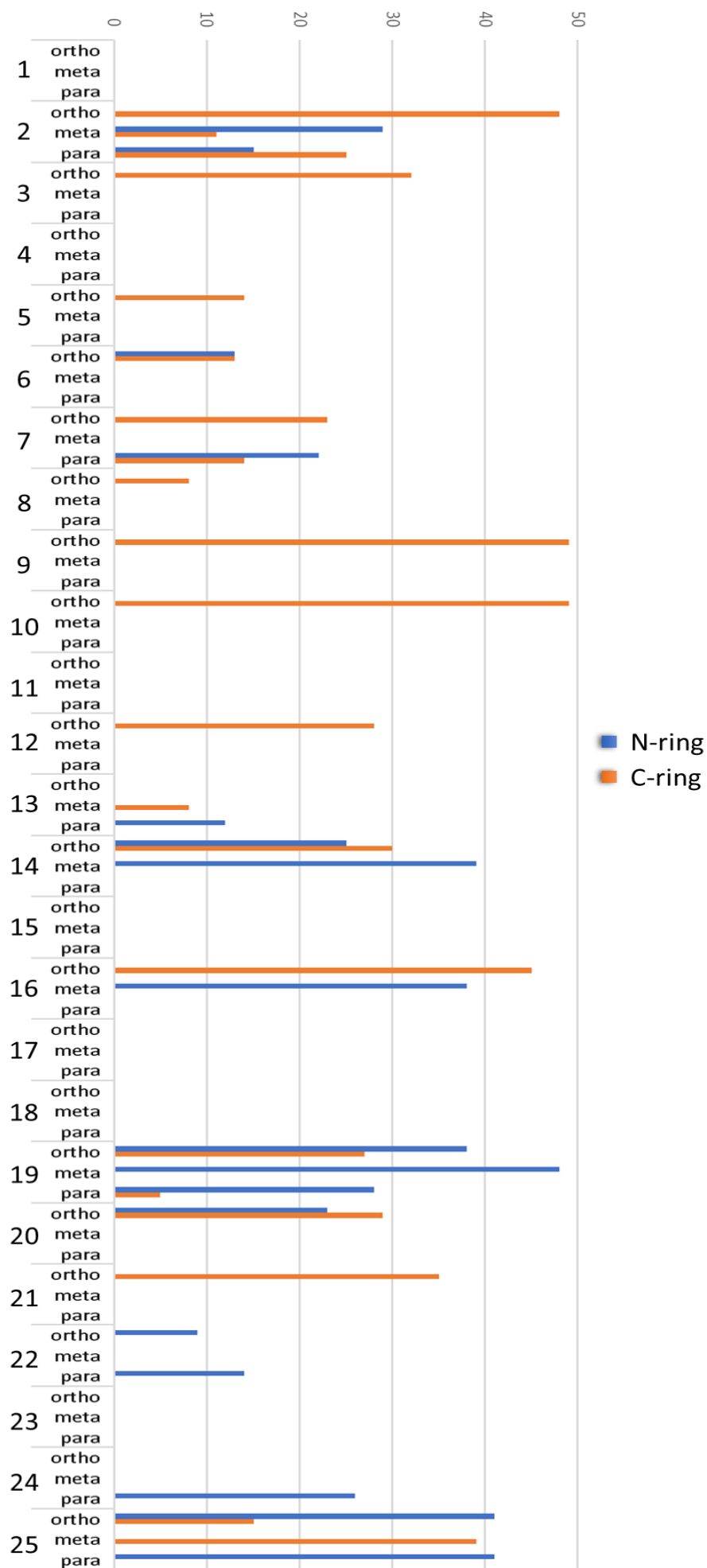
**Figure 56:** Model of ligands from the library previously created (**R** in **Figure 52**), in complex with the iridium transition metal.

#### 3.3.5.4 Induced Fit Docking (IFD) of metal complexes

At first, Glide XP docking was attempted on the selected ligands including the metal complex, using the previous parameters and grid setup. However, the space in the pocket was too small to allow any valid poses by this method.

IFD calculations were then conducted. This technique gave flexibility in the protein scaffold by refinement of the residues around the ligand (see details in section 3.2.2, page 62). The previous Glide docking grid was used to run the IFD calculations. From the 53 poses entered, 45 poses were recovered after IFD calculation among which 34 structures had only one aromatic ring connected to the metal support. Only few poses were obtained for the ligand with two aromatic rings. The structures were too elongated and complicated to be synthesised and none of them showed a top score. Based on these results, only the structures with a single aromatic ring were considered for the rest of the project.

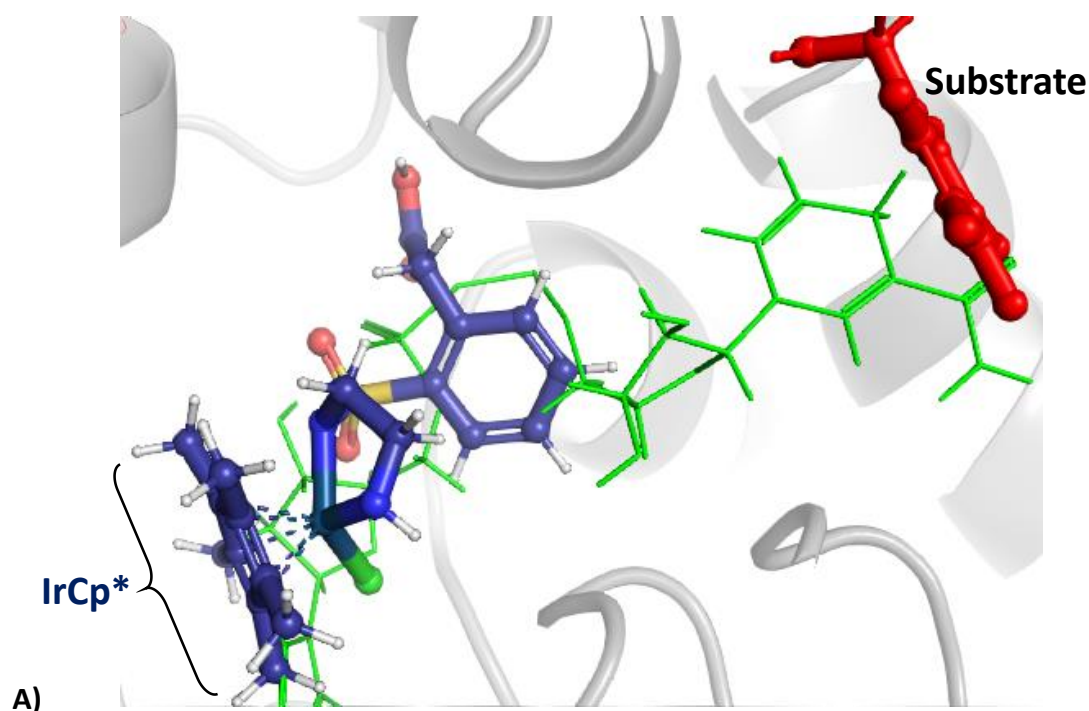
The poses, direction, orientation, and affinity scores were treated following the filters created previously (section 3.3.5.1, page 88). **Figure 57** presents the results of the filters applied to the 34 ligand poses from IFD calculations.

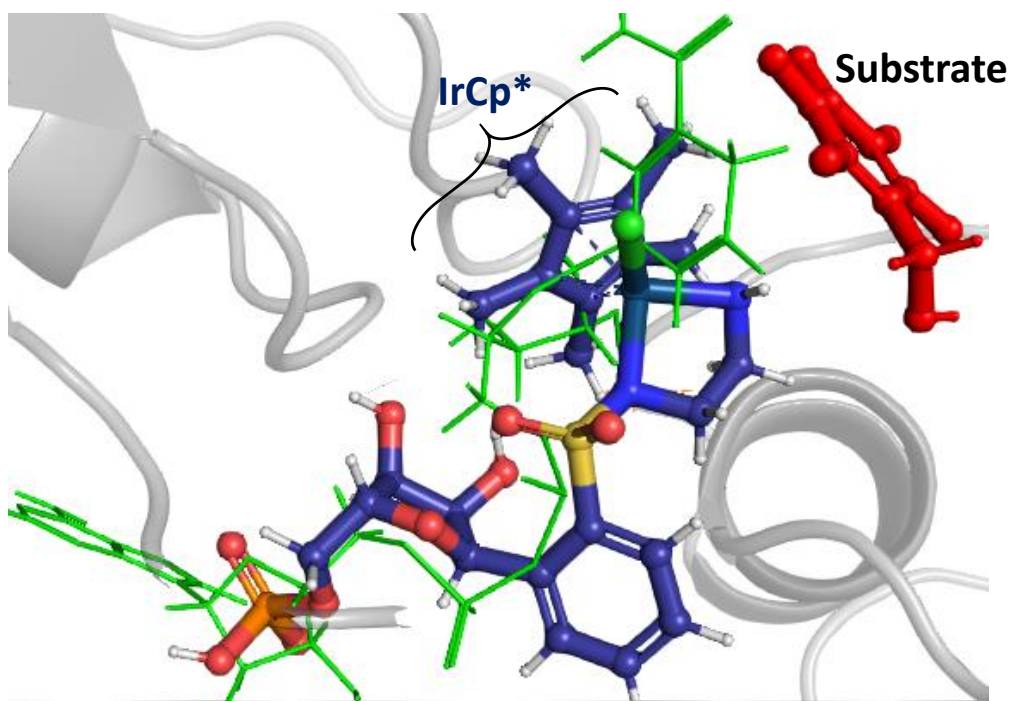


**Figure 57:** Graph representing the final score (y axis) obtained from the affinity score and the structures poses score (orientation/location) after HLADH docking. All the designed ligands (x axis) are represented with various metal complex insertions (ortho, meta, para) and the C-ring mode (orange) or the N-ring (blue).

The analysis of the obtained scores gave a quick overview for the docking process conclusions. From view of the substitution position at the aromatic rings no para substituted ligand reached an advantageous high score, while the number of highly scoring phenyl-based ring structures was similar to pyridine counterparts. This indicates no discrimination between the two in terms of binding affinity. However, it is relevant to highlight that one ligand structure was best ranked for only one type of ring. For example, ligand **R9** and **R10** obtained only best scored poses for the C ring, whereas if no impact of the ring occurs, both N and C ring structures should show the same score. Finally, the ranking of ligand structures depended mainly on the nature of the **R** substituent.

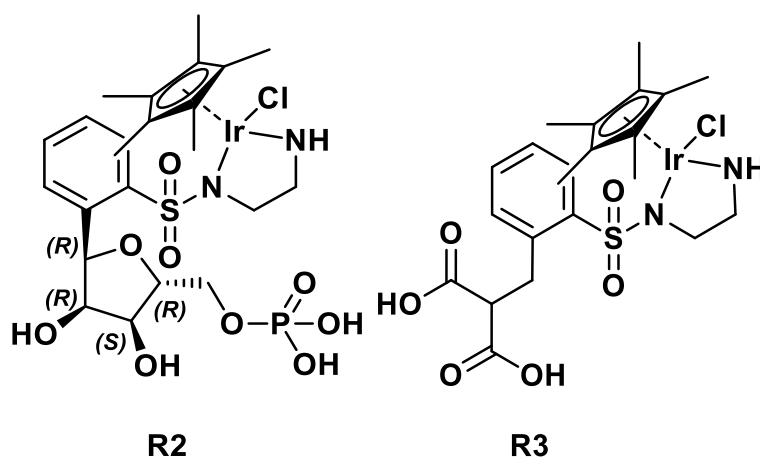
The applied filters naturally discredited wrong poses, for example structures with the metal opposite to the substrate or with few contacts. **Figure 58** depicts the structure of ligands **R2** and **R5**, which ranked with highest and lowest score, respectively.



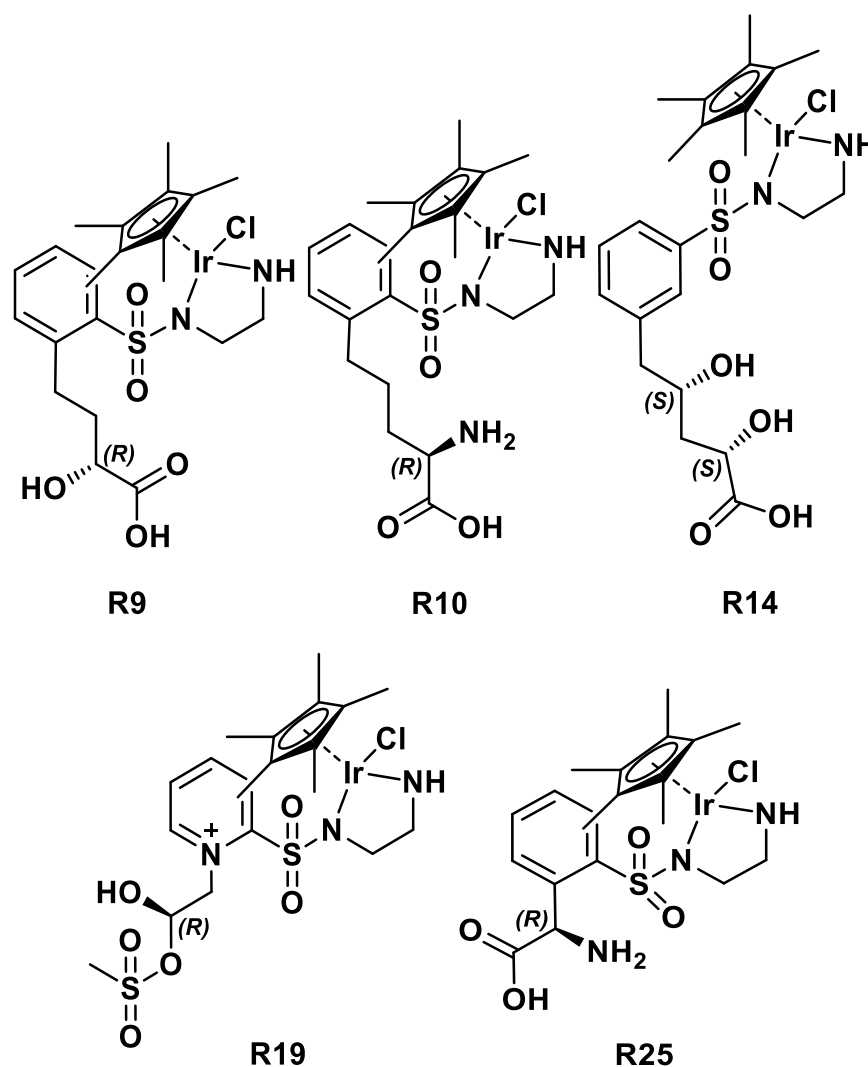


**Figure 58:** A) Ligands **R5<sub>ortho</sub>** (XP= -4.0 kcal mol<sup>-1</sup>, blue) with a wrong orientation for catalysis (superposed to NADH green). B) Ligand **R2<sub>ortho</sub>** (XP= -6.1 kcal mol<sup>-1</sup>, blue) with a suitable pose toward the substrate (red).

From the second docking screening, several best scoring ligands were selected: **R2**, **R3**, **R9**, **R10**, **R14**, **R19** and **R25** (Figure 59).







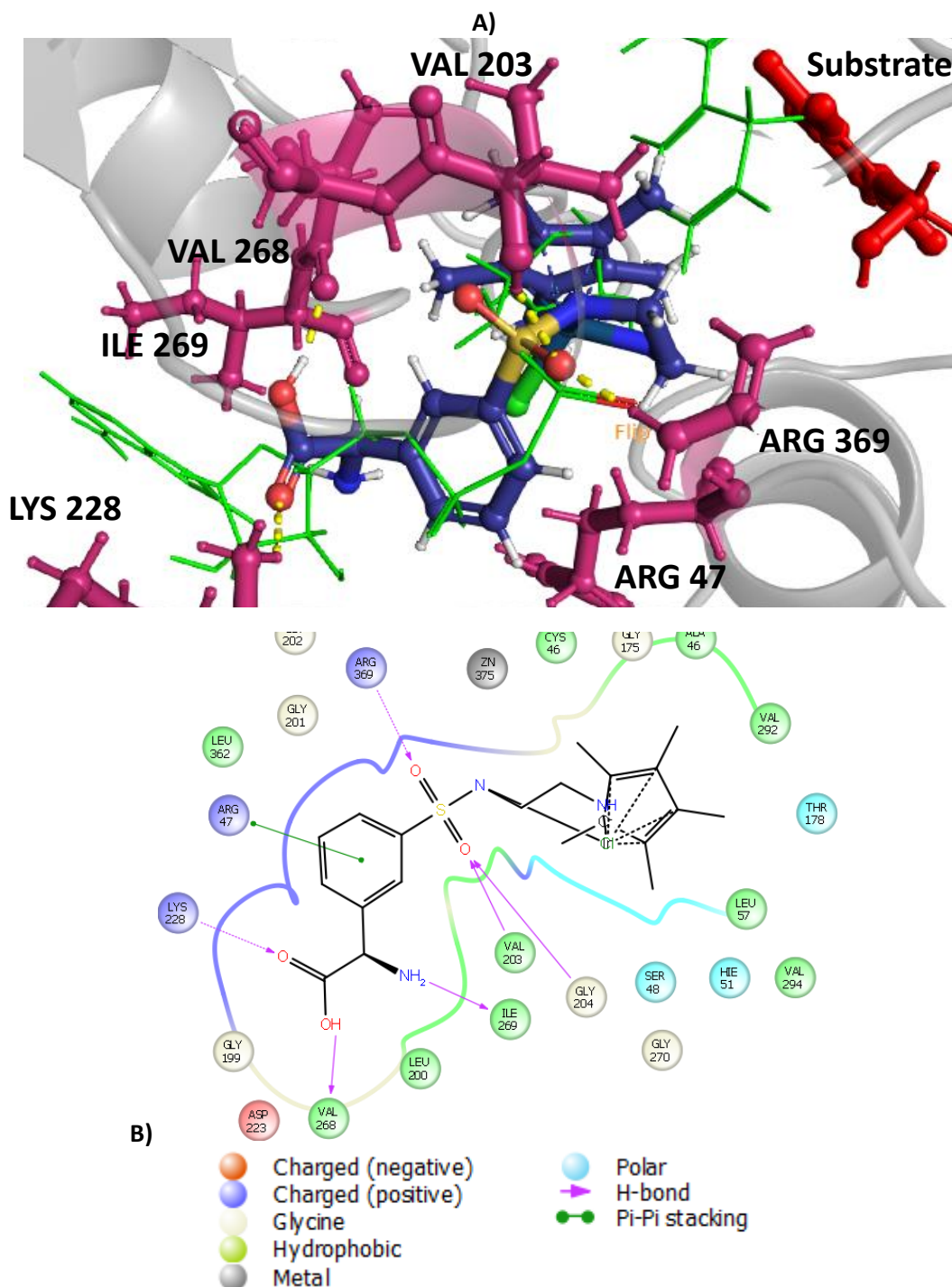
**Figure 59:** Set of lead catalyst structures, selected based on the IFD results. The anchoring part of these products should be suitable for the anchoring of the iridium transition metal inside HLADH.

Among these structures, the anchoring part of **R2** represented the nicotinamide mononucleotide part of NADH cofactor. The presence of this ligand was not surprising and confirms the finding of the initial docking, which highlighted ribose and phosphate as essential for the binding affinity.

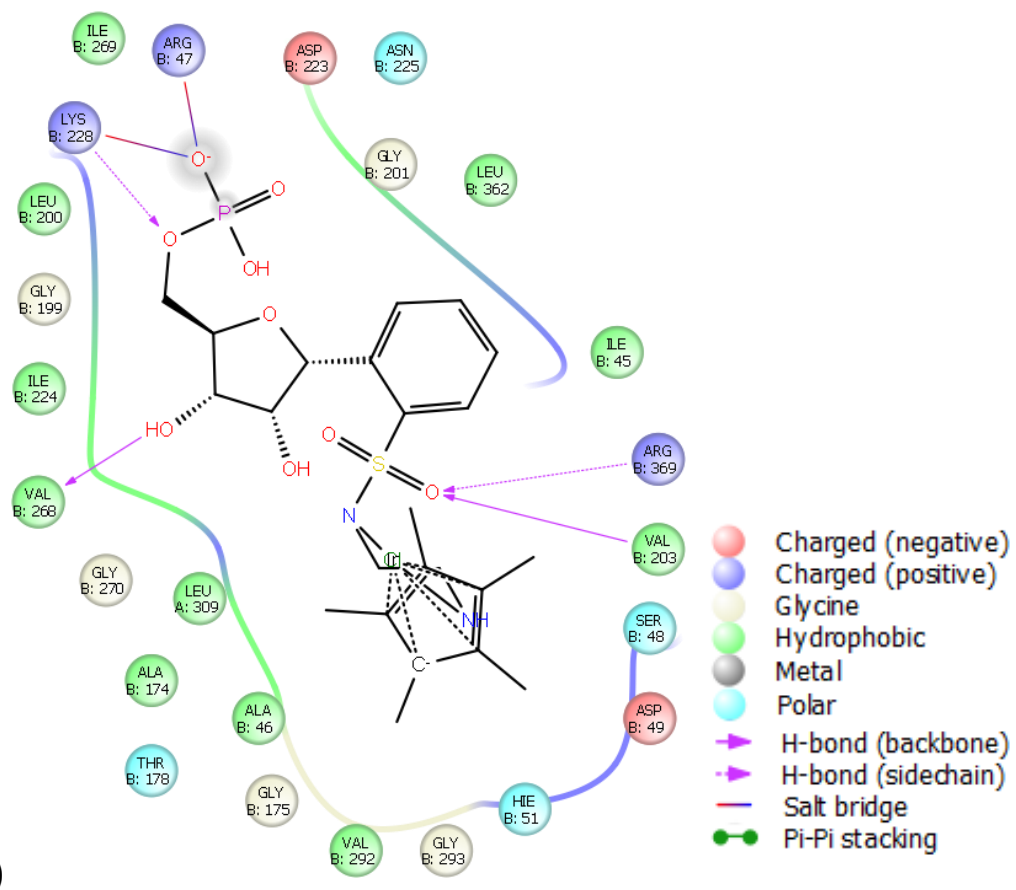
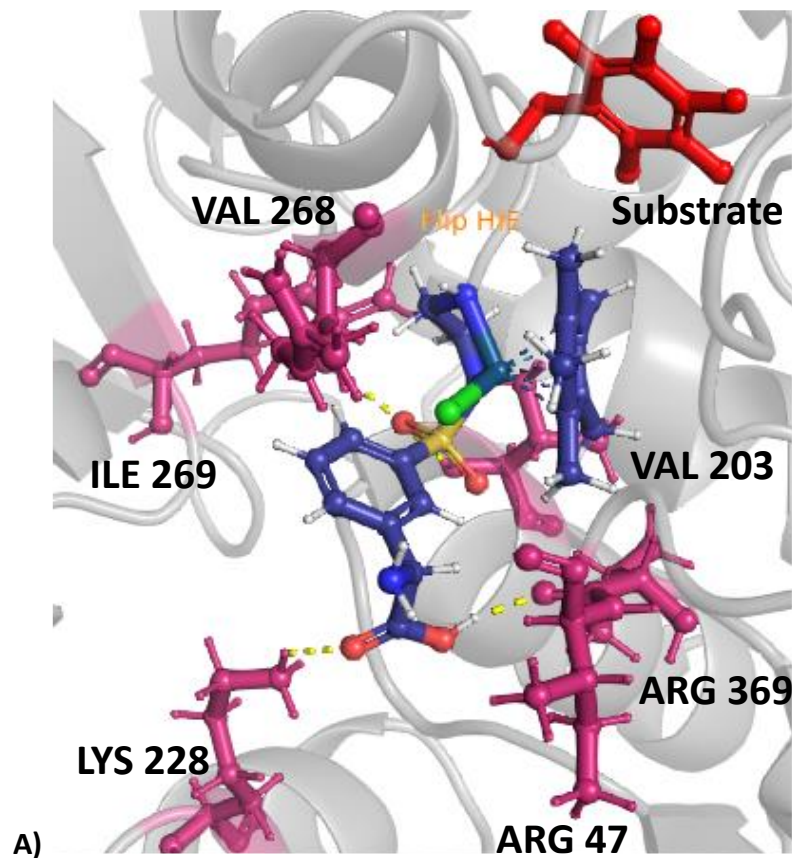
In other hit structures, hydroxyl and carboxylic acid groups appeared as a recurrent substituent, found in almost all structures. These observations matched with the first aim, the substitution of ribose (by hydroxyl) and phosphate (by carboxylic acid).

Conversely, ligands without carboxylic acid group, including **R1** or **R18** (**Figure 53**), presented no available pose. Ligands with only one functional group (e.g. ligand **R5**) were ranked with a weak score and wrong poses (**Figure 58**). The interaction diagram

and 3D poses of the hits ligands confirmed previous conclusions on the contacts with HLADH scaffold (examples in **Figure 60**; **Figure 61**; **Figure 62**):

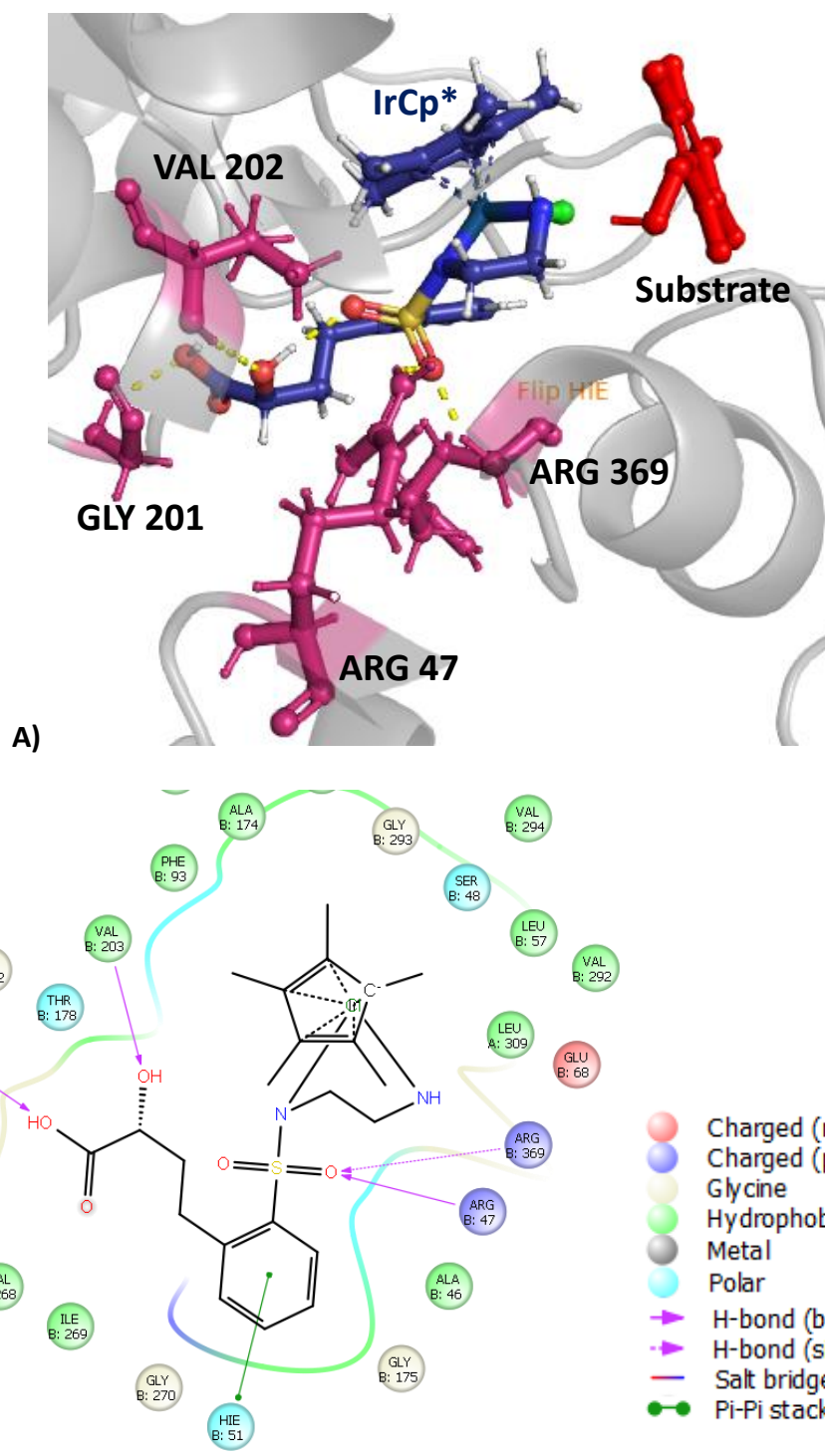


**Figure 60:** Ligand **R25**, with a C-ring and a meta insertion of the metal complex. **A)** Superposition of **R25** (blue) with **NAD<sup>+</sup>** (green). **B)** Interaction diagram of **R25** inside HLADH cofactor pocket. Arg: arginine, Asp: aspartic acid, Cys: cysteine, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Ser: serine, Thr: threonine, Val: valine.



**Figure 61:** Ligand **R2**, with a C-ring and a meta insertion of the metal complex. **A)** Superposition of **R2** (blue) with **NAD<sup>+</sup>** (green). **B)** Interaction diagram of **R2** inside HLADH cofactor pocket. Arg: arginine,

Asn: Asparagine, Asp: aspartic acid, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Lys: lysine, Ser: serine, Thr: threonine, Val: valine.



**Figure 62:** Ligand **R9**, with a C-ring and an ortho insertion of the metal complex. **A)** superposition of **R9** (blue) with **NAD<sup>+</sup>** (green). **B)** Interaction diagram of **R9** inside HLADH cofactor pocket. Arg: arginine, Ala: alanine, Glu: glutamic acid, Gly: glycine, His: histidine, Ile: isoleucine, Leu: leucine, Phe: phenylalanine, Ser: serine, Thr: threonine, Val: valine.

Several interactions were observed between the sulphonamide group and the important arginine residues. Additionally, new contacts with other residues were also spotted: K228, H51, S48 or various glycines.

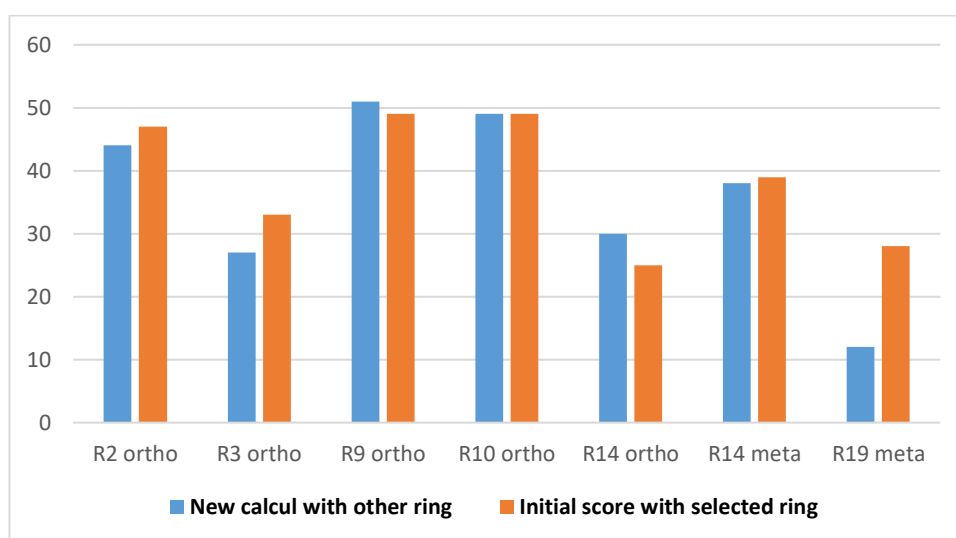
### 3.3.5.5 Validation of the process and hit selection

Following the selection of the best residues, a validation of the protocol was performed to avoid any false positive or negative results. In the first section, the docking protocol was already validated by using the RMSD method (section 3.3.1.1, page 63). This part focuses on the accuracy of the selection method applied.

#### - *N/C- ring differing scores*

As mentioned earlier, several ligands scored highly only with one type of the aromatic ring, obtaining no or poor results with the other ring structure (**Figure 57**). However, similar poses detected for C- and N- ring structures suggested that the nicotinamide ring showed no real impact in the affinity for HLADH pocket.

In order to understand these differences between the two types of rings and validate the process, a new set of IFD calculations with similar filters were carried out with two structures containing the two ring types (**Figure 63**).



**Figure 63:** Comparison of IFD scores between the same catalyst structures with either a N or C-ring. The structures used for the new IFD experiment showed scores similar with both rings, confirming that the nature of the ring had no effects in affinity. Thus, to understand the low scores in the original calculations, the whole process containing XP and SP docking selections was analysed in detail.

In the first Glide filtering, many selected structures (e.g. ligand **10**) appeared with only one type of ring. The other structures were generally ranked too low to be

considered in the 80 first selected structures. The reason behind this lies in the representative XP scores lying so close together that some structures with similar but slightly lower scoring fell below the cut-off for selection at this step.

Another explanation involved the conformation of the structures. Some ligands (e.g. **R2** or **R9**) were selected with both N and C ring after the SP / XP docking. However, a superposition between the two structures showed a difference in the conformation. During IFD calculations, a rigid treatment of the ligand was kept (while allowing flexibility for the enzyme) and the previous selected conformations stayed unchanged inside the pocket. In the end, the filters resulted in a bad scoring of the conformation for one of the structures. Thus, the differences in scores were due to different ligand conformers selected in the previous step, and not the nature of the ring.

In order to verify these hypotheses, the IFD approach was repeated for several other ligands structure, randomly checked for their differences between C and N ring IFD scores. All of them resulted in similar scores for both ring types when the same ligand conformation was used in the IFD input.

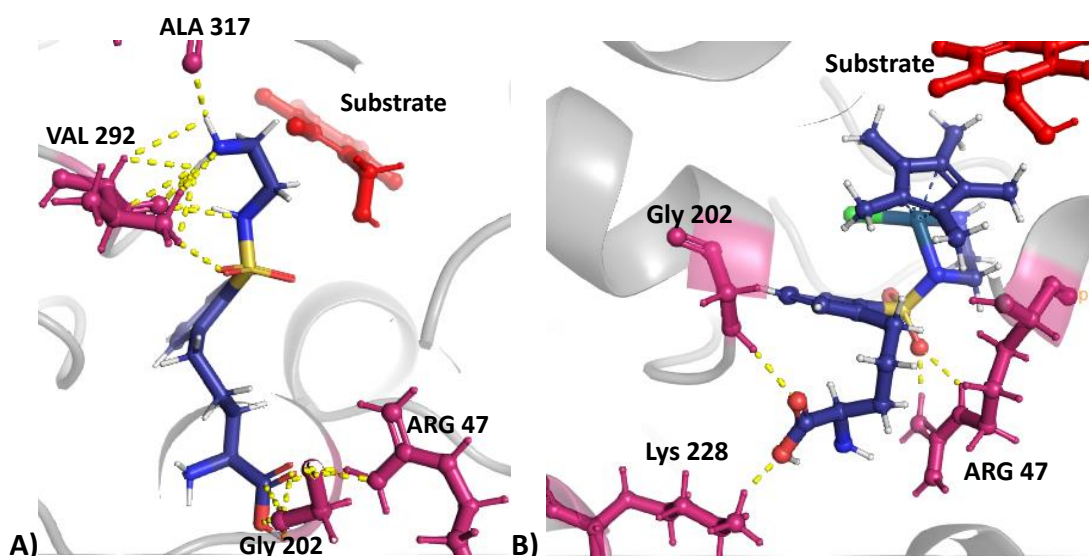
To conclude on this protocol validation, the nature of the ring had no impact on the binding process and the structure selection.

- ***Comparison of IFD and previous scores***

The best scores obtained by the selected metal lead structures were then compared with several of the previous Glide XP results.

- IFD result (metal ligands) vs Glide XP docking (only sulphonamide anchor):  
The Glide XP scores achieved by the docking of anchors are slightly higher than the IFD scores obtained for the corresponding metal ligands. For example, ligand **R10** scored -10.3 kcal mol<sup>-1</sup> without the metal complex (XP docking) while ligand **R10** = -7.0 kcal mol<sup>-1</sup> with the IrCp\* (IFD calculation). This score variation highlighted a more important number of contacts made with surrounding residues (**Figure 64**). Especially, the sulphonamide and amine functions were free during XP docking, allowing hydrogen bond to be created, contrary to IFD experiments where they were engaged in IrCp\* complex.

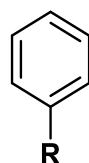




**Figure 64:** The number of contacts compared between **A)** an anchors **R10** with only sulphonamide support and **B)** metal complex catalyst.

- IFD results for metal catalysts against NADH cofactors:  
The NADH cofactor scored always better than XP or IFD docking of the catalysts, predicting a better affinity of the natural cofactor for HLADH.
- IFD results of metal catalysts against benzylnicotinamide ligands (**R**= Benzyl):  
The benzylnicotinamide anchor structure previously used were metalated to match the structure of the hit ligand. Thus, both N and C ring were used, and a sulphonamide support was added to the aromatic ring. The best Hollmann ligand obtained a low score of  $-4.5 \text{ kcal mol}^{-1}$ , with all structures in reversed position. These results were lower than the lead ligands (around  $-9 \text{ kcal mol}^{-1}$ ), meaning a poorer affinity.
- **Enhanced affinity tests of hits by fragment docking**

Fragments of the hit structures were tested through the second docking protocol and compared to the cofactors NAD(P)H (**Figure 50, Table 7**). Only the ring and the functional group **R** were employed to determine the impact of the anchoring part.



**Figure 65:** Truncated hit ligands used for enhanced affinity tests. They were composed only of an aromatic ring and the **R** anchoring part.

This test gave another indication on the affinity strength of the substituents.

**Table 7:** Glide XP results (kcal mol<sup>-1</sup>) of fragment into HLADH, and their position compared to the natural cofactor.

Ligands	XP scores	Pose compared to NAD <sup>+</sup>
NADH (1)	-14.3	-
NADPH (3)	-9.7	-
NAD <sup>+</sup> (2)	-13.3	-
NADP <sup>+</sup> (4)	-14.5	-
<b>1</b>	-3.5	reversed structure/ no contact
<b>R2</b>	-8.6	similar location / contacts
<b>R14</b>	-8.2	similar location /contacts similar to ribose and pyrophosphate of NAD <sup>+</sup>
<b>R9</b>	-7.4	same as <b>R14</b>
<b>R10</b>	-6.6	same as <b>R14</b>
<b>R25</b>	-6	same as <b>R14</b>
<b>R19</b>	-5.3	same as <b>R14</b>

The hit structures were not at the level of the natural cofactors due to missing hydrogen bonding interactions. Yet, as a shorter structure, the truncated hit structure still obtained acceptable scores, between -8 and -5 kcal mol<sup>-1</sup>. Furthermore, they clearly showed better scores than the benzylnicotinamide ligands **1**, which experimentally obtained low affinity for HLADH.<sup>106</sup> Inhibition assays from M. Basle's thesis also concluded to a lack of interactions with the HLADH for all the different benzyl ligand.<sup>109</sup>

The selected ligand presented lower affinity than the natural NADH cofactor. Nevertheless, they were the best structures tested by docking, with significant higher affinity compared to existing NADH mimics from literature.

### 3.3.6 Comparative study with TbADH

#### 3.3.6.1 Introduction

Before proceeding with the synthesis of the lead anchor structures, another alcohol dehydrogenase (ADH) enzyme was tested.

The *Thermoanaerobacter brockii* alcohol dehydrogenase (TbADH) is an NADP-dependent, zinc ADH from bacteria, currently studied in the laboratory for covalent



anchoring.<sup>97</sup> The accessibility to this enzyme, a known higher thermostability than HLADH, the resistance to organic solvent and a broad substrate specificity made TbADH another interesting target for the engineering of AMs.<sup>101</sup>

Anchor structures previously designed were also tested inside TbADH cofactor binding pocket to evaluate prospective affinities. This enzyme structure could also become an alternative at HLADH scaffold if no efficient experimental results were found.

### 3.3.6.2 Comparison of docking with HLADH

#### - *Cofactor docking*

The TbADH crystal structure was without substrate (PDB 1YKF).<sup>97</sup> The docking setup was exactly similar to the one used earlier for HLADH. Re-docking of crystal ligand NADP<sup>+</sup> gave an RMSD < 2 Å, validating the protocol for this enzyme.

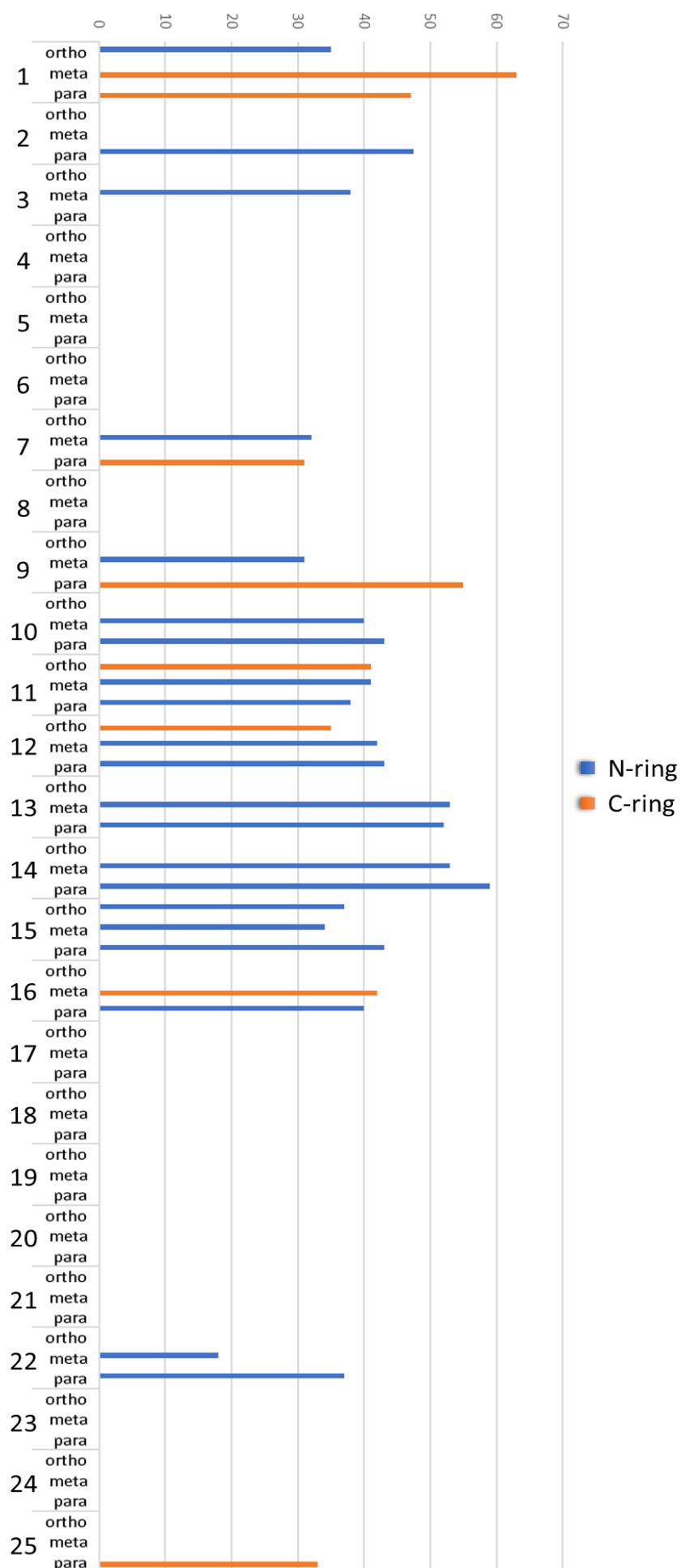
**Table 8:** Docking of NAD(P) cofactors in TbADH, following a docking protocol set in the previous section (3.3.1.1, page 63). The Glide XP scores are in kcal mol<sup>-1</sup>.

Enzyme	NADH	NAD <sup>+</sup>	NADPH	NADP <sup>+</sup>
TbADH	-11.2	-8.0	-11.6	-11.7
TbADH no Zn <sup>2+</sup>	-10.5	-8.7	-12.3	-11.1

Similar to HLADH docking experiment, the Glide scores were not able to clearly discriminate between NADH and NADPH cofactors, except for a lower affinity with NAD<sup>+</sup> (**Table 8**). TbADH sequence identity is only about 27 % with HLADH,<sup>97</sup> with the greatest similarity in the cofactor binding pocket. However, the literature suggests a specificity for NADPH favoured by the adenine binding to three amino acids (S199, R200 and Y218). In HLADH, this space is occupied by an aspartic acid (D223), repelling the phosphate.<sup>215</sup>

#### - *Library docking*

The anchor library designed for HLADH was docked inside TbADH with the previous protocol starting with Glide docking and followed by IFD calculations for the metal complexes (**Figure 66**). The TbADH crystal structure was prepared like HLADH, the zinc ion was removed and all three amino acids around (C37, H59 and D150) were mutated to alanine.



**Figure 66:** Graph representing the final score (y axis) obtained from the affinity score and the structures poses score (orientation/location) by IFD calculations in TbADH. All the designed ligands (x axis) are represented with various metal complex insertions (ortho, meta, para) and the C-ring mode (orange) or the N-ring (blue).

In TbADH, more structures were ranked with a better score after application of the filters than previously in HLADH. This could be explained by the presence of a substrate in HLADH, reducing the space for Cp\* group compared to TbADH free of substrate. Thus, the limit was raised to a score of 40 for ligand to be selected as a hit (against 30 for HLADH). A broader range of structures in the right catalytic location and orientation were selected as hits: **R1, R2, R9, R10, R11, R12, R13, R14, R15, R16** (red structures have also been selected in HLADH docking).

This table compares the scores and ranking of TbADH and HLADH for their hit anchors:

**Table 9:** Comparison of hit anchors in TbADH and both HLADH with and without substrate. The ranking of the structures is made by their filters score. XP score ( $\text{kcal mol}^{-1}$ ) obtained from IFD calculations is also shown.

Hit structures	TbADH XP score	Hit structures	HLADH no substrate XP score	Hit structures	HLADH XP score
<b>R1</b>	-14.0	<b>R16</b>	-10.4	<b>R9</b>	-7.0
<b>R14</b>	-11.7	<b>R13</b>	-8.8	<b>R10</b>	-7.2
<b>R9</b>	-10.3	<b>R16</b>	-9.4	<b>R16</b>	-7.3
<b>R13</b>	-11.2	<b>R3</b>	-7.8	<b>R2</b>	-6.1
<b>R2</b>	-12.0	<b>R14</b>	-9.6	<b>R19</b>	-6.2
<b>R16</b>	-10.3	<b>R6</b>	-9.2	<b>R14</b>	-8.6
<b>R10</b>	-10.0	<b>R12</b>	-8.0	<b>R25</b>	-5.1
<b>R12</b>	-10.0	<b>R10</b>	-8.4	<b>R21</b>	-5.0
<b>R15</b>	-11.0				
<b>R11</b>	-11.8				

The two studies in HLADH showed lower affinity scores for the anchors than TbADH. The affinity was even lower in the substrate presence, maybe again due to less space for Cp\* or because of amino acids conformation in the “close mode”, refraining the

binding of ligands. Indeed, according to literature, more amino acid movements are observed in HLADH scaffold upon the binding of cofactors. The scaffold conformation “close” around the cofactor, which can explain the difficulties of binding for Cp\* (especially in crystal structure with substrate).<sup>91</sup> In TbADH, the main conformation changes are localised in the adenine binding area for the specificity for NADP cofactor.<sup>97</sup>

In conclusion, TbADH seemed to be less discriminating, accepting a broader range of ligands. Within the first nine hit structures, TbADH and HLADH without substrate shared identical ligands (**Table 9**). Besides, four ligands were common to the three enzymes and part of the hit anchors previously selected: **R10**, **R14** and **R16**. The visual analysis in HLADH showed similar H-bonds made by these anchors and the co-crystallised cofactor. The conclusion of docking in TbADH again highlight that all the previously mentioned catalysts (**R10**, **R14** and **R16**) made similar contacts with TbADH scaffold than NADP<sup>+</sup>.

With the selected hit anchors also in TbADH docking top ranking, this enzyme can be another scaffold to use for AMs engineering. Therefore, the synthesised anchors will be tested into both HLADH and TbADH (available in the laboratory) to compare their affinity for the enzymes and validate the docking protocol.

### 3.4 Conclusion and selection of hit structures

Molecular docking was used as a tool to design metal complexes, able to strongly bind inside HLADH. A first docking protocol was developed and validated for the classical determination of the binding affinity of natural NAD(P)H cofactors, previous NAD(P)H mimics and ligands libraries. Various enzyme residues and functional groups were then highlighted as necessary for a strong affinity. These findings enabled the creation of a NAD(P)H analogue model and the design of a small library of potential metal catalyst. A second docking protocol was developed, which analysed these new metal catalysts for their affinity, but more importantly, for their orientation and location toward an optimal catalysis. It is decisive to present the metal complex towards the substrate pocket, but without blocking the substrate binding site.

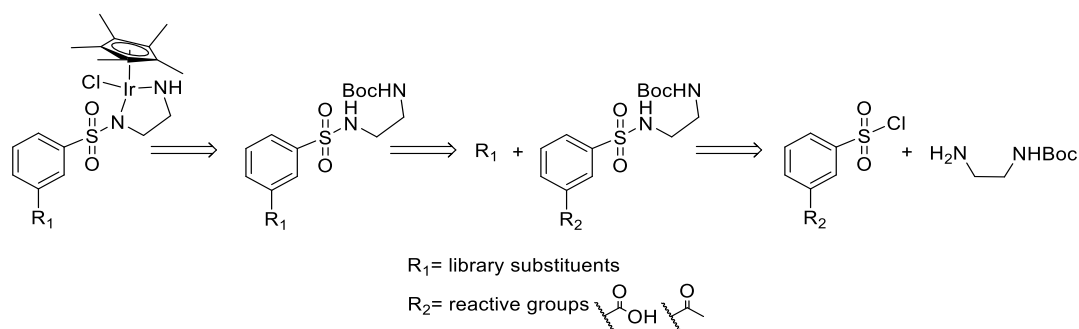
This new process allowed the selection of a set of metal complexes ranked with the best docking and location score. The affinity approximation of the metal catalysts was lower than the natural cofactor. However, the selected metal catalyst ranked better than benzyl nicotinamide ligands, already tested experimentally with a low affinity for HLADH. Finally, these catalysts were also tested in TbADH and reported similar results than HLADH docking. TbADH was also selected as another potential scaffold for AM. Amongst the selected lead catalysts, some with the best scores in both enzymes (**R2**, **R9**, **R10** and **R14**) are already favoured for the next synthesis step.

## 4 Synthesis of metal-based catalysts for the supramolecular anchoring of iridium transition metal catalysts inside ADHs

### 4.1 Background

The previously designed and computationally tested metal catalysts were composed of three main parts. First, a *N*-sulfonyl-ethylenediamine moiety, derived from the Noyori's catalysts, coordinates the transition metal.<sup>56</sup> Several studies have shown that *N*-sulfonyl-ethylenediamine complex of Ru and Ir were the most efficient catalysts for the aqueous asymmetric transfer hydrogenation of imines.<sup>83,118,123</sup> The second part of the catalysts consisted of the functionalised aromatic ring, acting as an anchor to the nicotinamide-binding pocket of alcohol dehydrogenases (ADHs). From the docking calculations, there were no significant differences in the affinity to ADHs between compounds bearing pyridine and phenyl rings. Thus, the phenyl ring was preferred for an easier and quicker synthesis. The third part of the catalysts was a supramolecular ligand, with features to bind inside ADHs natural cofactor binding pocket with high affinity (e.g. hydroxyl, carbonyl). Following the computational work, lead structures of ligands were selected based on their docking scores and their optimal binding orientation. The meta-position of *N*-sulfonyl-ethylenediamine moiety was preferred to mimic the position of the amide in the natural NADH cofactor. A multistep retrosynthetic workflow was then designed toward the production of these metal catalysts (**Scheme 1**).

The retrosynthesis proposed was developed to fulfil two parameters: 1) a divergent synthesis promoting the creation of a ligand library through the functionalisation of a single intermediate 2) a short synthesis, yielding the metal catalysts in few steps.



**Scheme 1:** Retrosynthesis designed to produce the previously selected hit metal catalysts structures. The starting structure was a tosyl ring holding a functional group (**Scheme 1**,  $R_2$ ). The first step was to protect the *N*-sulfonyl-ethylenediamine by a *tert*-butyloxycarbonyl protecting group (Boc), a common protection for amines. The Boc protecting group can be cleaved under strong acidic conditions and resists to basic conditions and nucleophilic attacks. From previous protocols with similar *N*-sulfonyl-ethylenediamine, this protection has been used for further modifications of the aromatic ring without impact on the amine.<sup>57,220</sup>

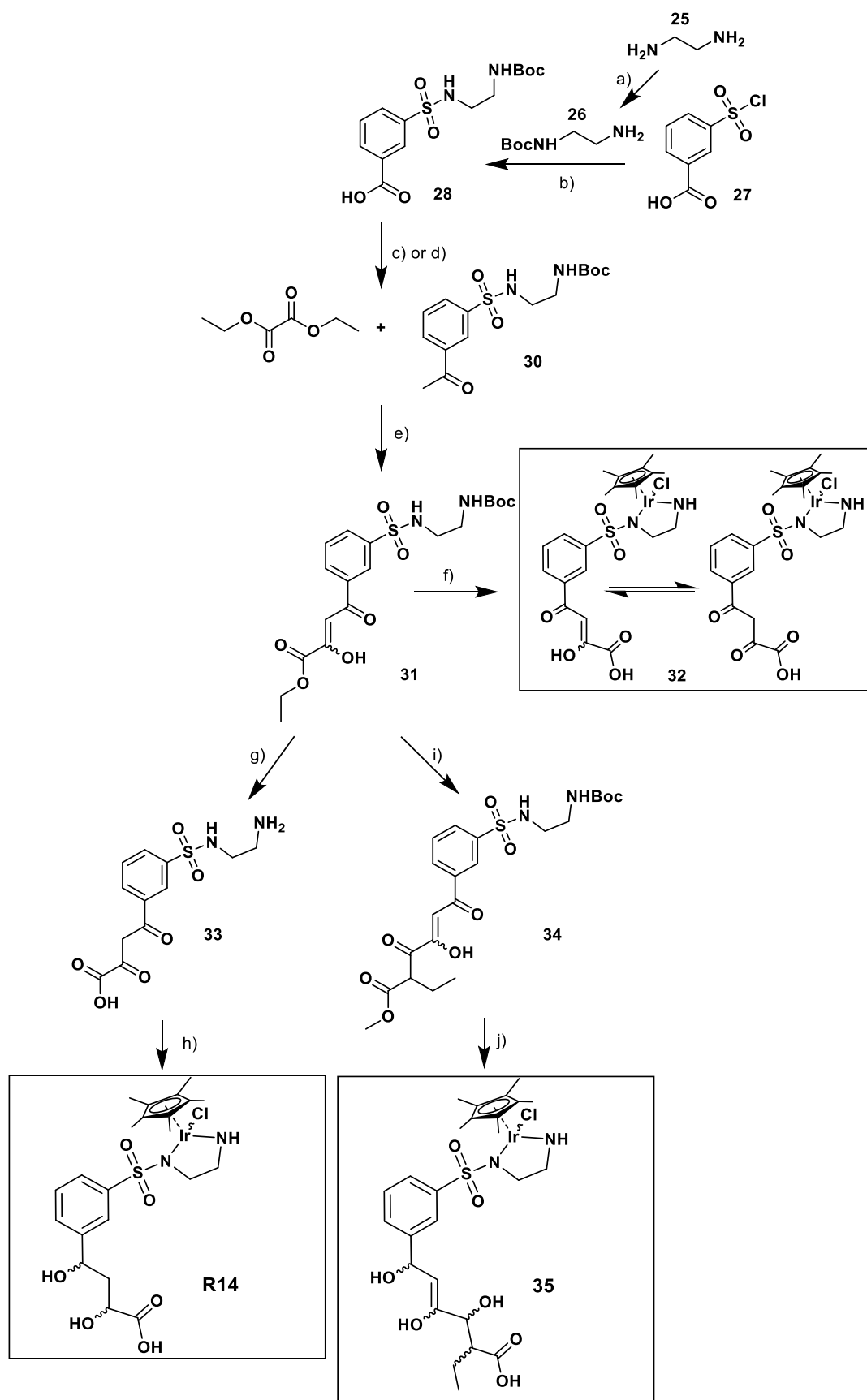
The second step was the development of a small library of structures by substitution or addition of new functional groups (**Scheme 1**,  $R_1$ ). Finally, a metalation reaction was promoted to obtain an iridium metal catalyst for supramolecular anchoring into ADHs.

A divergent synthetic pathway was then designed. The main objective was to synthesise a *N*-sulfonyl-ethylenediamine intermediate containing a reactive functional group able to be modified into different ligands. Based on the structures suggested by the docking studies (**R2**, **R9**, **R10**, **R14**), three metal catalyst complexes were targeted: **R14**, **32** and **35** (**Scheme 2**). The number and diversity of hydrophilic functional groups and the anchor's length was an attractive parameter for future affinity investigations. Compound **R14** was thus privileged over **R9** and **R10** because its variety of carbonyl and hydroxyl group combinations, to mimic the binding of the ribose and phosphate of NADH. The presence of several hydrogen bond donor and acceptor groups made these ligands potential better supramolecular anchors compared to phenyl substituted compounds studied previously.<sup>108</sup> Besides, unlike complicated structures (e.g. the NMN ligand **R2**, **Figure 59**),<sup>221</sup> the selected ligands

allowed for a short synthesis route of 4 - 6 steps, consuming less time and being cost effective.

The designed synthetic pathway was based on the functionalisation of a the *tert*-butyl (2-((3-acetylphenyl)sulfonamido)ethyl)carbamate intermediate **30**, through a Claisen condensation, to form the key ethyl 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate **31**. After applying subsequent Claisen condensation<sup>222</sup> or reduction reaction, this key structure can be converted into the different target molecules:





**Scheme 2:** General synthetic pathway for the production of metal catalysts. Reagents and conditions: a) Boc<sub>2</sub>O, DCM; b) triethylamine, Boc-ethylenediamine, DCM; c) (i) SOCl<sub>2</sub>, CH<sub>3</sub>NHOCH<sub>3</sub>, Et(NPr-*i*)<sub>2</sub>, dry DCM; (ii) CH<sub>3</sub>MgBr, dry THF; d) CH<sub>3</sub>Li, DEE; e) *t*-BuOLi, dry THF; f) (i) TFA, DCM, (ii) NaOH, MeOH/H<sub>2</sub>O, (iii) [IrCl<sub>2</sub>Cp\*]<sub>2</sub>, Et<sub>3</sub>N, MeOH; g) (i) NaOH, MeOH/H<sub>2</sub>O, (ii) TFA, DCM h) (i) NaBH<sub>4</sub>, MeOH, (ii) [IrCl<sub>2</sub>Cp\*]<sub>2</sub>,

Et<sub>3</sub>N, MeOH; i) *t*-BuOLi, ethyl propionate, dry THF; j) (i) NaBH<sub>4</sub>, MeOH, (ii) TFA, DCM, (iii) [IrCl<sub>2</sub>Cp\*]<sub>2</sub>, Et<sub>3</sub>N, MeOH.

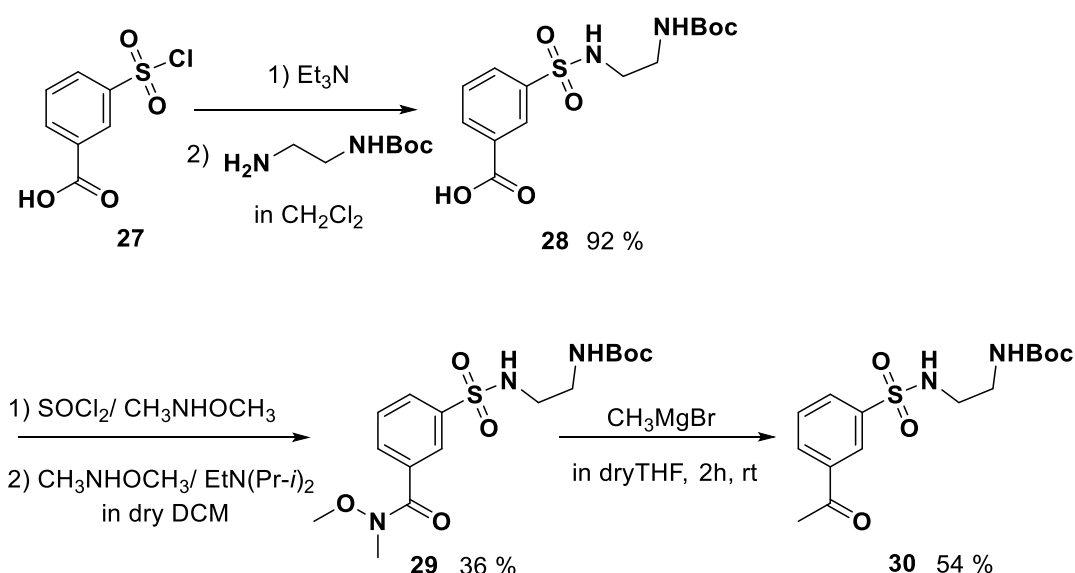
## 4.2 Results and discussion

### 4.2.1 Synthesis of the ketone-substituted intermediate

The first steps of the synthetic pathway were focused on the preparation of the ketone intermediate **30**, from which a Claisen reaction led to the key compound **31**, a common product to the different metal catalyst synthesis.

Commercial chlorosulfonyl benzoic acid **27** was used as starting material, because no chlorosulfonyl acetophenone was commercially available (**Scheme 3**). Compound **27** was treated with a protected *N*-Boc ethylenediamine in the presence of triethylamine, according to several literature protocols.<sup>220,223,224</sup> This step functionalised the aromatic ring with the *N*-sulfonyl-ethylenediamine protected by *tert*-Butyloxycarbonyl (Boc). The choice of a Boc protection allowed to keep intact the future metal coordinating group. The planned steps involved basic conditions (Claisen reactions), and no strong acidic conditions, which are known to remove the Boc protection.

Product **28** was obtained in 92 % yield after purification. This compound **28** was part of the designed ligands in the previous docking study (**R23**, **Figure 52**). The structure ended ranked with less affinity than the selected hits anchors. However, to validate the docking method, the affinity of **28** anchoring part can be tested against future compounds.

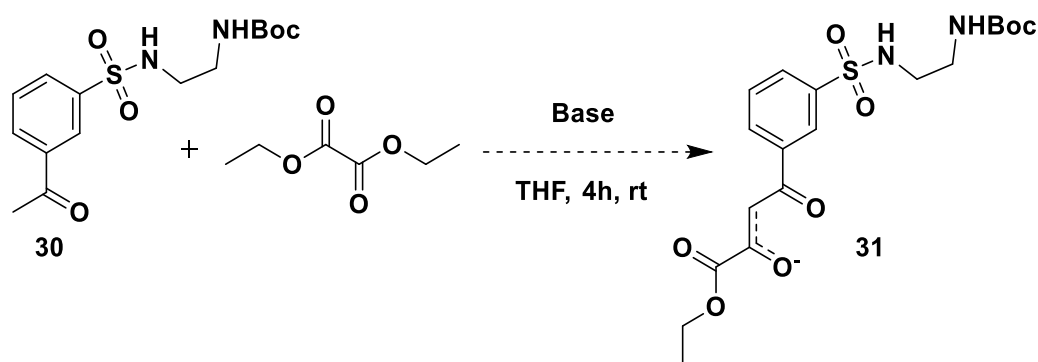


**Scheme 3:** The 3 steps synthesis of ketone **30**.

From the benzoic acid **28**, the ketone intermediate **30** was synthesised in two steps according to the Weinreb-Nahm ketone method.<sup>225,226</sup> Based on this method, the benzoic acid **28** was first activated by conversion into a Weinreb-Nahm amide **29**, with a 36 % yield. Subsequent treatment of the amide **29** with a commercial Grignard reagent led to the ketone **30**. The final step was performed in a dry solvent under inert atmosphere, to bypass possible reaction of the methyl magnesium bromide solution with water. The pure intermediate **30** was obtained with a 54 % yield from the *N*-Boc Weinreb amide **29**. The starting material amide was also recovered from purification with 14 % yield. The two steps overall yield was of 20 %. The limiting step was the activation of the acid in the Weinreb amide. In literature, the resulting amide was not purified, so the amide **29** was then used in the next step without purification (3 g of acid **27** gave 1.7 g of ketone **30**).

#### 4.2.2 Synthesis of the key intermediate $\beta$ -diketoester

Compound **31** represented a challenge to succeed in the route to the future metal catalysts library. The synthesis of **31** was performed by Claisen condensation of diethyl oxalate with ketone **30** to give a  $\beta$ -diketoester functional group (**Scheme 4**).<sup>227</sup> The general conditions involved the use of a base to create an enolate anion, which would perform a nucleophilic attack on the carbonyl functionality of the ester.



**Scheme 4:** Claisen condensation scheme of ketone **30** and diethyl oxalate, for the synthesis of product **31**.

Diethyl oxalate was chosen to react with the ketone **30** to achieve the targeted carbonyl pattern. A range of experiments was performed with variations of the conditions adapted from literature (**Table 10**).

In an initial attempt, the sodium ethoxide (NaOEt) was selected as a base, according to previous work on diketoesters synthesis.<sup>227,228</sup> Following the established protocol,

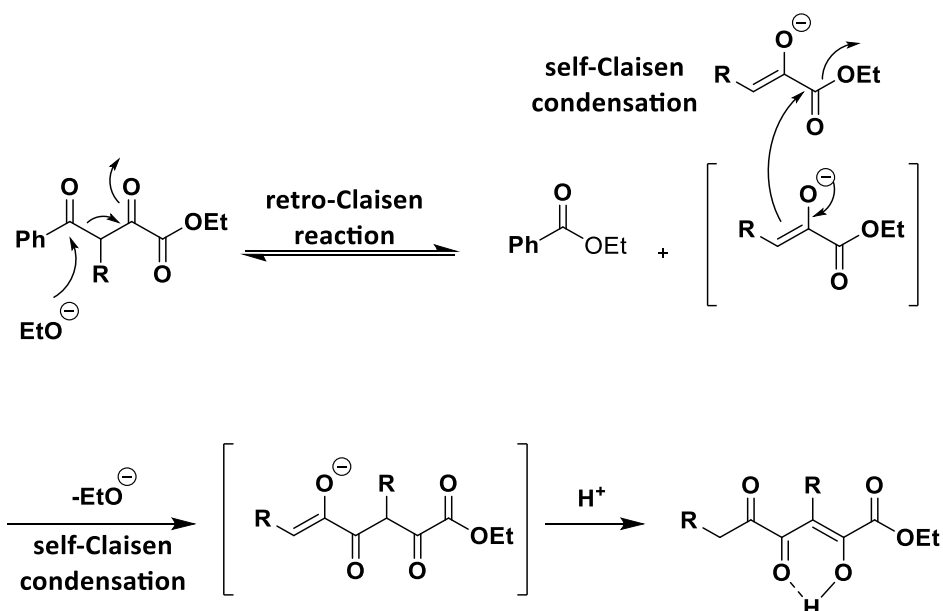
NaOEt was first utilised at 1 eq (**1**, **Table 10**). The analysis by thin layer chromatography (TLC) showed only the presence of the starting material after 16 hours of reaction. Verification of the  $^1\text{H-NMR}$  spectrum confirmed that the mixture was composed of the starting material, with presence of signals attesting for only few, uninterpretable by-products. No changes were observed after increasing the NaOEt concentration up to 4 eq.<sup>227</sup> Thus, according to other literature examples,<sup>229</sup> the solvent was changed to control its impact on the reaction (**3**, **Table 10**). The starting material and the previously described by-product were again detected in the  $^1\text{H-NMR}$  spectrum, with high amount of a new by-product with signals at 4.39 ppm and 1.40 ppm. Unfortunately, the reaction did not provide the expected  $\beta$ -diketoester **31**.

It was decided to change the base to sodium hydride (NaH), which is commonly used for Claisen rearrangements (**4**, **Table 10**).<sup>230,231,232</sup> NaH concentration was raised until disappearance of the starting material on TLC analysis. After study of the  $^1\text{H-NMR}$ , the expected product was detected in low yield (12 %), with a mixture of by-products and starting material. Optimisation was needed to obtain the desired compound **31** with a relevant yield.

**Table 10:** Claisen condensation reaction between ketone **30** and diethyl oxalate under various conditions.

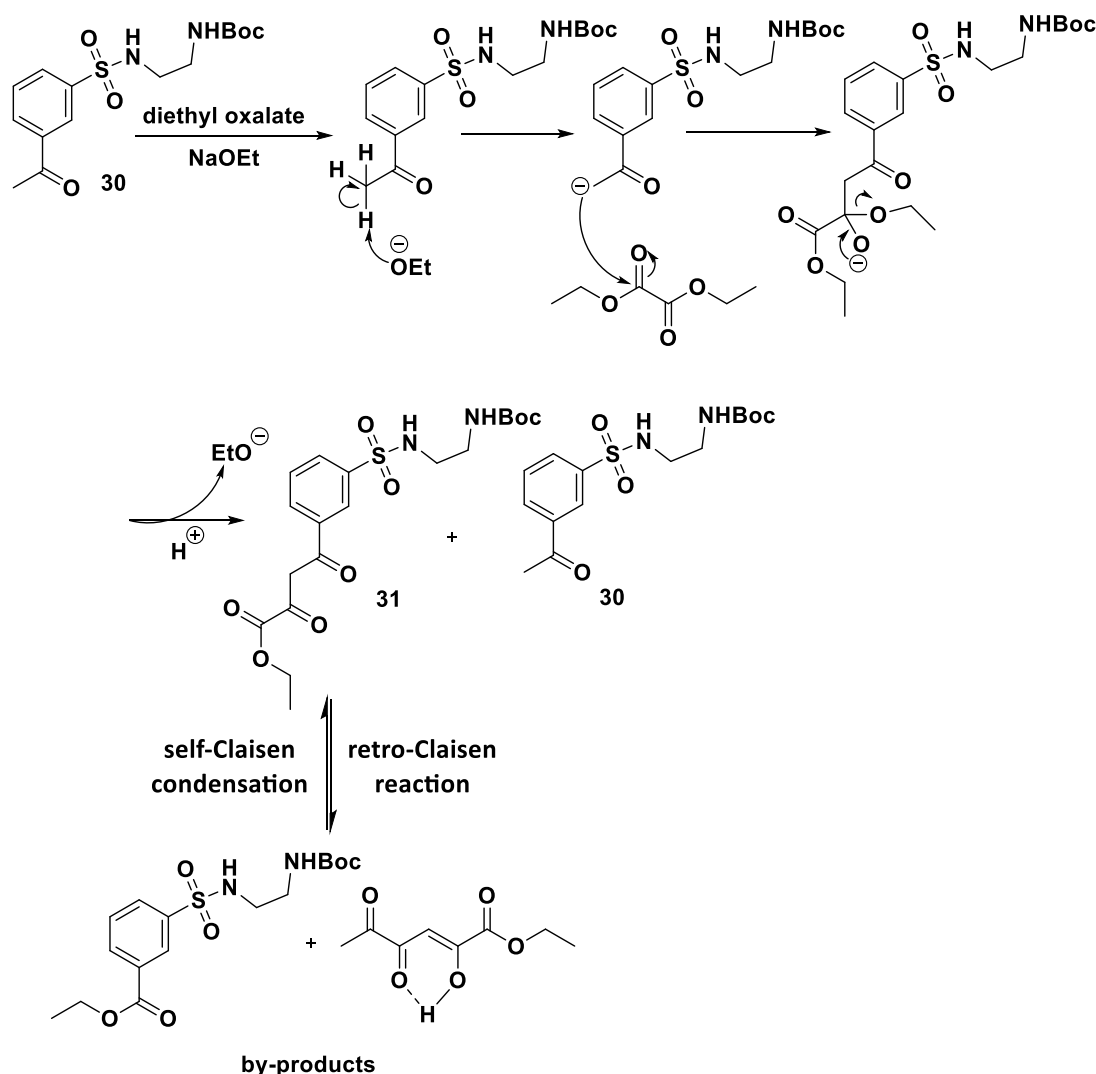
Entry	Base	Eq.	Solvent	Reaction time	Outcomes
<b>1</b>	NaOEt	1	Et <sub>2</sub> O	Overnight	Starting material + by-products
<b>2</b>	NaOEt	4	Et <sub>2</sub> O	Overnight	Starting material + by-products
<b>3</b>	NaOEt	2	Methanol	Overnight	Starting material + by-products
<b>4</b>	NaH	5	Et <sub>2</sub> O	Overnight	Product 31 + by-product and starting material

A report by Ji and co-workers<sup>233</sup> demonstrated a low yield of their expected  $\beta$ -diketoester when following the NaOEt-mediated sterically hindered Claisen condensation (**Scheme 5**). The authors explained that by-products observed in their  $^1\text{H-NMR}$  spectra came from a retro-Claisen condensation and a self-Claisen condensation:



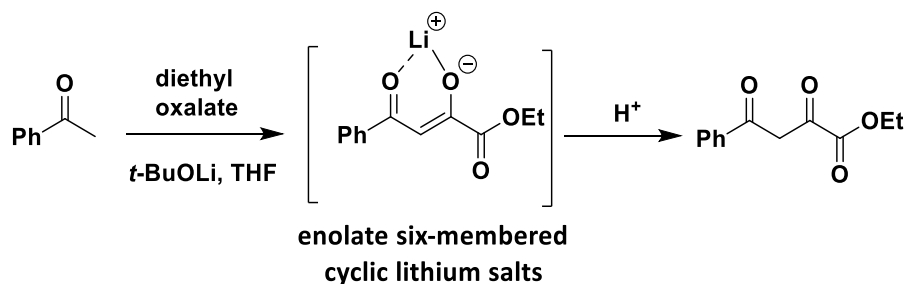
**Scheme 5:** Representation of a retro-Claisen reaction and a self-Claisen condensation mechanism, adapted from the proposed mechanism by Ji and co-workers.<sup>233</sup>

The by-products  $^1\text{H}$ -NMR spectra, described in their paper, correspond to the by-product signals found in the performed NaOEt attempts of Claisen condensation (**1** to **3**, **Table 10**). When using NaOEt or NaH to condense ketone **30** and diethyl oxalate, the formation of  $\beta$ -diketoester was followed by additional reactions generating the observed by-products, without possible isolation of the  $\beta$ -diketoester **31** (**Scheme 6**).



**Scheme 6:** NaOEt catalysed Claisen condensation between ketone **30** and diethyl oxalate. Representation of the by-products created by self-Claisen condensation and retro-Claisen reaction, based on Ji and co-workers proposed mechanism (**Scheme 5**).<sup>233</sup>

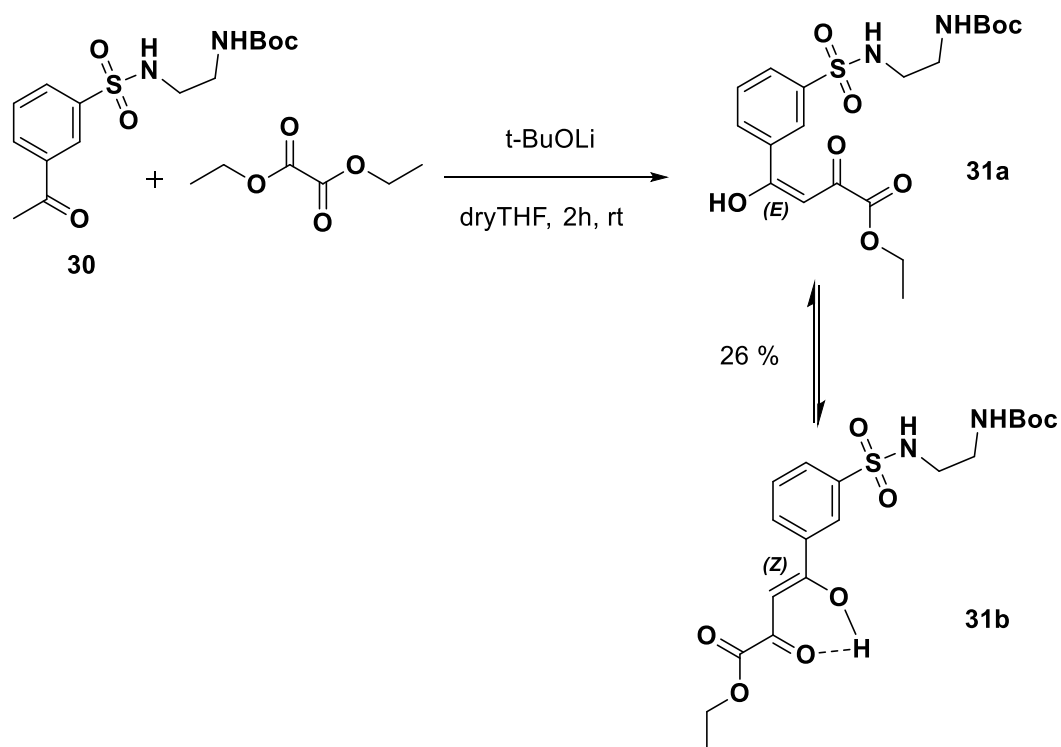
Ji and co-workers suggested the base plays an important role in the stabilisation of the reaction intermediates (**Scheme 7**). By creating a stable six-membered enolate, the base should avoid the side reactions previously mentioned.



**Scheme 7:** Example of the stabilisation mechanism of *t*-BuOLi mediated Claisen condensation, adapted from suggestions by Ji and co-workers.<sup>233</sup>

Thus, the authors proposed to use lithium bases instead of sodium-based ones. Due to the strong affinity of lithium for oxygen, the base was able to form a cyclic enolate

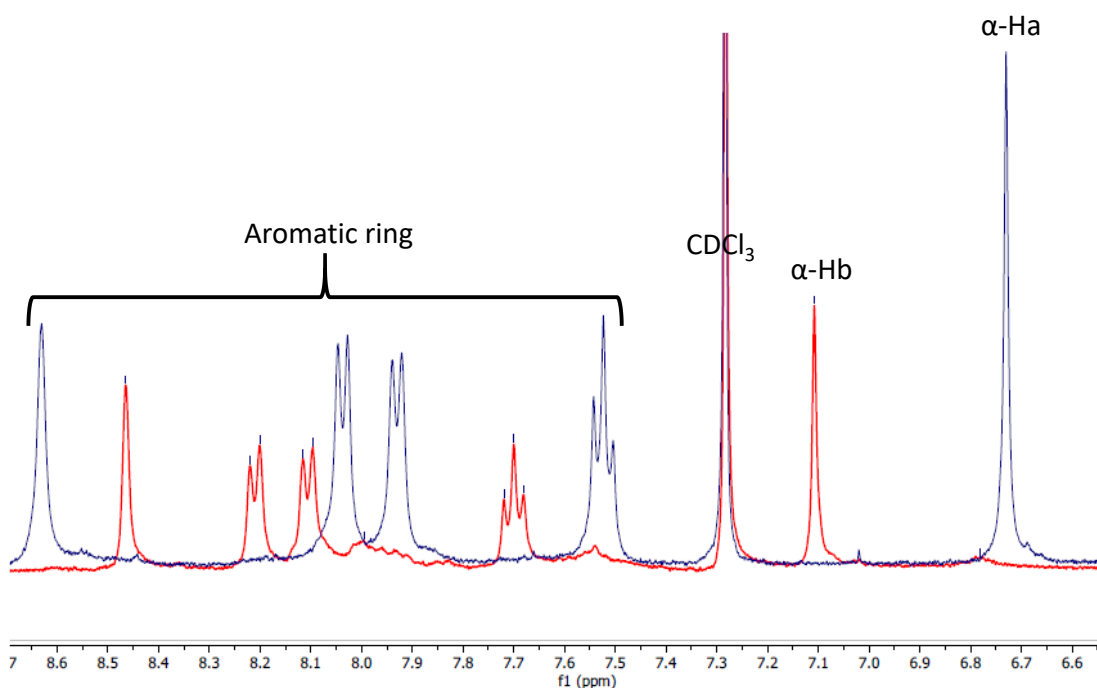
lithium salt, which circumvented second reactions. Their protocol used the *tert*-butoxide (*t*-BuOLi) base (**Scheme 8**). Following a similar protocol, the  $\beta$ -diketoester **31** was obtained with an overall 26 % yield.



**Scheme 8:** Adapted protocol for the synthesis of product **31** by Claisen condensation.

TLC analysis showed two distinct spots, corresponding to the two enolic forms *trans* and *cis* of the  $\beta$ -diketoester **31**, later purified by column chromatography (86 % *cis*-**31b** and 14 % *trans*-**31a** mass). The *cis*-enol was the most present and the most stable form due to the six-membered ring formation with intramolecular hydrogen bond.<sup>234,235</sup>

The  $^1\text{H}$ -NMR spectrum of both compounds **31a** and **31b** showed an integration of only one hydrogen for the peak at the vinyl position, confirming the presence of the enol form over the keto form of the  $\beta$ -diketoester (see Appendix 1). A different shift of the vinyl hydrogen and the aromatic ring signals was observed between the two spectra, which helped in confirming the *cis* and *trans* assignment (**Figure 67**). The vinyl proton in *cis*-**31b** showed higher signal frequency (7.1 ppm) than the *trans*-**31a** signal (6.7 ppm), due to the proximity of a carbonyl. The deshielding effect of the carbonyl was also observed for the aromatic proton singlet, which shifted to higher frequency in the *trans*-**31a** spectrum compared to the *cis*-**31b**.



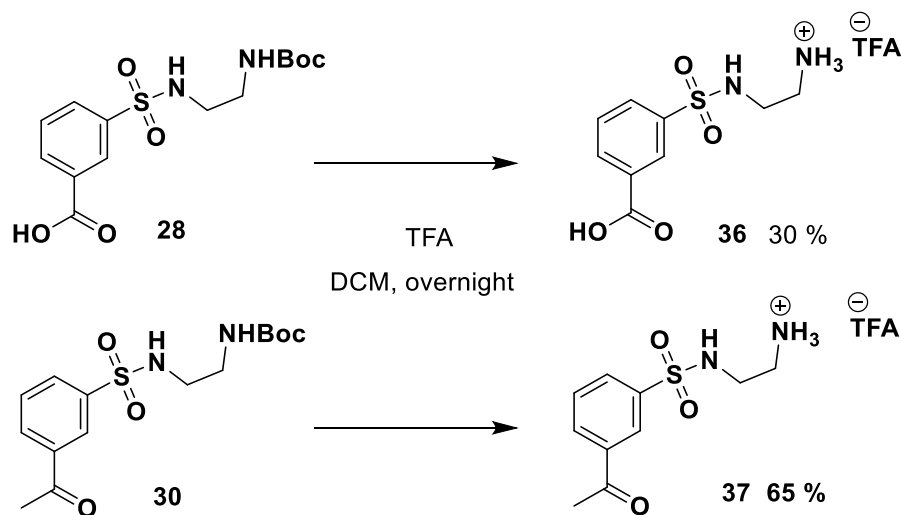
**Figure 67:**  $^1\text{H}$ -NMR spectrum comparison showing the proton shift differences between the two tautomers of the  $\beta$ -diketoester **31**: blue **31a** and red **31b** (full spectrum in Appendix 1).

These experimental results were consistent with literature observations. The  $\beta$ -dicarbonyls compounds exist in five tautomers: the ketone form, two *cis*-enol forms (with the intramolecular hydrogen migrating between the carbonyls) and two *trans*-enol forms.<sup>234,235</sup> The enol-enol migration of the intramolecular hydrogen between the two oxygens was too fast to be detected by  $^1\text{H}$ -NMR. Besides, in tricarbonyl product, the keto form is usually not found.<sup>229,235,236</sup> Thus, only the enol form *cis* and *trans* were isolated. Both enol forms were used in the next step, the deprotection of *N*-sulfonyl-ethylenediamine.

#### 4.2.2.1 *N*-Boc deprotection of the $\beta$ -diketoester **31** and of its intermediates

The deprotection of the Boc protecting group was the final step before the insertion of the transition metal. Usually, a Boc deprotection uses strong acids, such as trifluoroacetic acid (TFA) or hydrochloric acid (HCl).<sup>220,237</sup> The protonation of the *tert*-butyl carbamate induces the loss of a *tert*-butyl cation. The HCl/TFA anion decarboxylate the carbamic acid left, to result in the free amine later protonated under acidic conditions.





**Scheme 9:** TFA catalysed deprotection of Boc intermediates **28** and **30** to afford the deprotected carboxylic acid **36** and ketone **37**.

The *N*-Boc intermediates acid **28** and ketone **30** were deprotected by an excess of TFA, affording the products **36** (30 % yield) and **37** (65 % yield). The  $^1\text{H}$ -NMR spectrum validated the results with absence of the  $\text{CH}_3$  signal (see Appendix 1). These two new products were then used in the next step of metalation to create new catalyst complexes.

To the best of our knowledge, the deprotection of Boc in a molecule with a 1,3- $\beta$ -diketoester has not been tested in the literature. Only one experiment was described in a patent from Boojamra and co-workers, using TFA reagent.<sup>238</sup> Other successful examples also promoted deprotection with TFA in DCM on molecules containing hydroxy ethyl ester, substituted ketoesters or substituted 3,5-diketoester molecules.<sup>239,240,241,242</sup> Thus, the TFA-mediated deprotection protocol used earlier was applied to the  $\beta$ -diketoester **31** (**Scheme 9**).

The concentration of TFA was varied following recommendations from the authors. According to TLC monitoring, the reaction showed no changes at low concentration and no starting material was left at high concentrations of TFA (**1**, **2**, **Table 11**). Unfortunately, the reaction did not proceed to the deprotection of Boc group. The  $^1\text{H}$ -NMR spectrum of the crude mixture from high concentrations presented similar proton signals as the deprotected ketone **37**, along with the degradation of *N*-sulfonyl-ethylenediamine by loss of the relevant  $^1\text{H}$ -NMR signals.

Based on several literature protocols, variations of conditions were attempted as displayed in **Table 11**.

**Table 11:** Different conditions applied for the Boc deprotection reaction of **31**.

Entry	Reagent	Eq.	Solvent	Reaction time	Outcome
1	TFA	12	DCM	Overnight	Degradation + Ketone <b>37</b>
2	TFA	8	DCM	Overnight	Starting material
3	HCl	1 M	Methanol, EtOAc, water, dioxane	Overnight	Starting material
4	HCl	4 M	EtOAc, methanol	Overnight	Ketone <b>37</b> + by-products + starting material
5	HCl	Conc.	Methanol	Overnight	Ketone <b>37</b> + by-products
6	H <sub>2</sub> SO <sub>4</sub>	1.5	Dioxane	6 h	Starting material
7	H <sub>3</sub> PO <sub>4</sub>	15	THF; DCM	1 day	Starting material
8	Iodine	1	DCM	Overnight	Starting material
9	-	-	Water reflux		Starting material

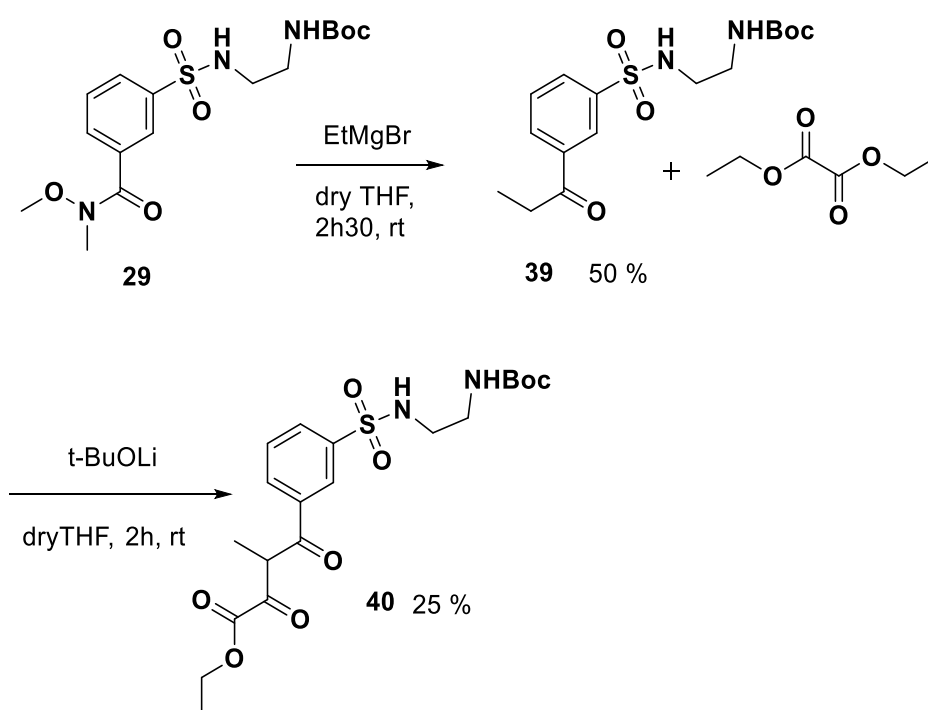
First, the acid reagent was replaced by HCl, the second most used acid in *N*-Boc deprotection. The concentration of acid and the solvent were varied based on several existing protocols (**3** and **5**, **Table 11**).<sup>220,243,244</sup> Similar to TFA results, reactions with HCl in low concentration ended with the starting material while concentrated HCl degraded the product to the deprotected ketone **37**. The <sup>1</sup>H-NMR and ESI-MS analysis of the crude mixture using 4 M HCl in methanol or ethyl acetate showed a mixture of products. The deprotected ketone **37** was once again the main by-product spotted.

To protect the sensitive part of the molecule, milder acidic conditions were tested. Various common protocols used H<sub>2</sub>SO<sub>4</sub> or H<sub>3</sub>PO<sub>4</sub> for successful *N*-Boc deprotection.<sup>237,245,246</sup> However, in both cases, the starting material was detected in the <sup>1</sup>H-NMR spectra. (**6**, **7**, **Table 11**). Literature protocols using non-acidic conditions were then investigated.<sup>246,247,248,249</sup> The use of aqueous reflux (100 °C) and of iodine proved no reactivity with the starting material observed by <sup>1</sup>H-NMR spectra (**8**, **9**, **Table 11**).

Because the reaction conditions were not suitable to the current structure, it was decided to modify the product **31**, while keeping an anchoring part able to bind

properly inside ADHs. Examples of a TFA catalysed reaction on diketones or ketoester structures substituted between the carbonyls by aryl, triphenylphosphine or carbonyl groups, were performed without changes in the diketone integrity.<sup>240,250,251</sup> In order to test the potential effect of a substituent between the two carbonyls, a methyl group was added between the two ketones (**Scheme 10**). The choice of a methyl group, instead of the phenyl group, was made after Ji and co-workers never succeeded to mediate the Claisen condensation with a phenyl substituted ketone as starting material.<sup>233</sup> Besides, the methyl also avoids the insertion of a large hydrophobic group, which might later impact on the affinity inside the enzymes cofactor pocket. As seen in the docking experiments, the anchor structure should increase its affinity when containing hydrophilic groups.

From the Weinreb amide **29**, the commercial Grignard reagent ethyl magnesium bromide was used instead of the previous methyl magnesium bromide.<sup>226</sup> The conditions were kept as mentioned before, to synthesised the expected ethyl ketone **39** with 50 % yield.



**Scheme 10:** The two steps synthetic pathway to produce the ethyl 4-(3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfonyl)phenyl)-3-methyl-2,4-dioxobutanoate **40**, following previously describe protocols.

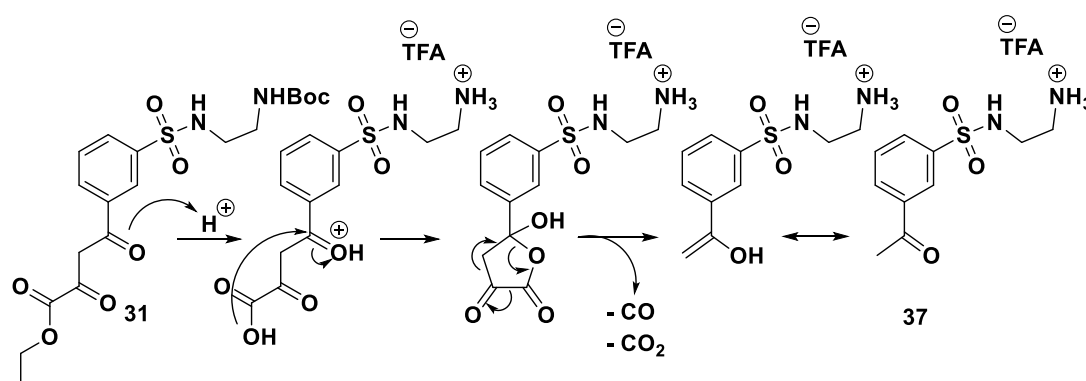
The previously optimised Claisen condensation conditions were used in the following step for the synthesis of the corresponding  $\beta$ -diketoester.<sup>233</sup> The methyl substituted

$\beta$ -diketoester **40** was obtained at 25 % yield, a similar yield than with the unsubstituted product.  $^1\text{H-NMR}$  spectrum displayed only the keto form over the enol, due to the presence of an alkyl substituent located between the two carbonyl groups (see Appendix 1).<sup>235,236</sup>

For the deprotection step, higher concentrations than the previous TFA and HCl-catalysed deprotections were needed before seeing the starting material consumed by TLC. Both resulting  $^1\text{H-NMR}$  spectra showed by-products mixture, with none of the expected  $\beta$ -diketoester signals. The by-products were difficult to determine with  $^1\text{H-NMR}$ , and only the deprotected ketone **37** was identified again. The methyl group helped stabilising the structure, however, not enough to deprotect the *N*-Boc without degradation of the  $\beta$ -diketoester. In order to further stabilise the  $\beta$ -diketoester part, a phenyl group would be an alternative substitution for a higher potential of hindrance effect. However, this group can also be challenging in the binding to the ADHs cofactor pocket.

After investigating on several reaction conditions and structure modifications, it was concluded that *N*-Boc deprotection of an aryl sulfonamide substituted with  $\beta$ -diketoester was not achievable. Only one patent apparently succeeded in TFA deprotection.<sup>238</sup>

Besides, Murray and co-workers proved degradation of a ketoester by TFA mediated decarboxylation.<sup>252</sup> Munro and co-workers then proved degradation of a natural diketoester into the corresponding ketone, by using low concentration of TFA (0.05 %).<sup>253,254</sup> Based on their decarbonylation-decarboxylation explanation, **Scheme 11** presents a hypothesis for the degradation mechanism of compound **31**.



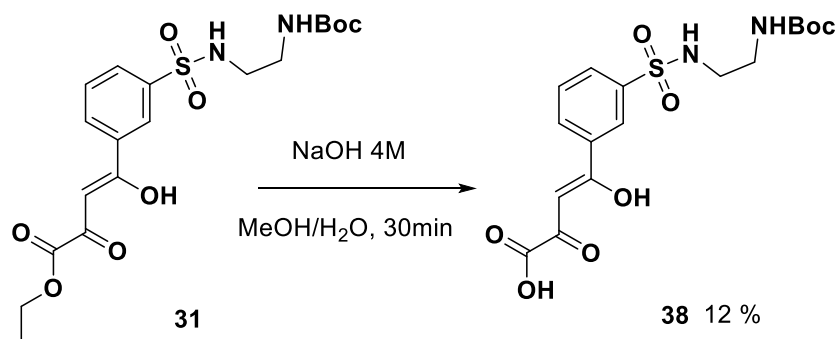
**Scheme 11:** Hypothesis of a degradation mechanism of compound **31** into the deprotected ketone **37**, following a decarbonylation-decarboxylation combination.

Under acidic conditions, the protonation of carbonyl starts the nucleophilic attack by the deprotected ester hydroxyl, forming a five-membered ring. A combination of decarbonylation and decarboxylation forms the ketone. Meanwhile, the acidic conditions also deprotected the *N*-Boc, affording the ketone **37**.

#### 4.2.2.2 Ester deprotection of the $\beta$ -diketoester **31**

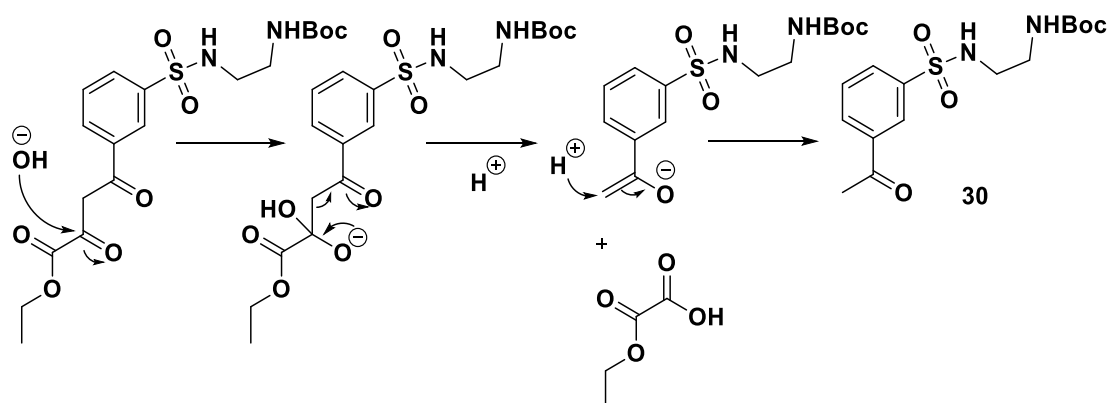
Given that the Boc protecting group was not removed without degradation of the  $\beta$ -diketoester, the focus was changed toward the ester deprotection instead. The size of Boc group is similar to the IrCp\* normally introduced at this position, therefore, affinity experiments inside ADHs were deemed to provide information about the metal complex and the anchoring part of the structure.

The saponification reaction was applied using basic conditions for the hydrolysis of the ester (thus avoiding any possible degradation of the  $\beta$ -diketoester from acidic conditions).<sup>255</sup>



**Scheme 12:** Ester deprotection of product **31** through saponification reaction to afford the product **38**. The hydrophilic anchoring part was composed of ketones and carboxylic acid.

The hydroxyl of NaOH works as a nucleophile and led, through ester deprotection, to the  $\beta$ -diketo acid **38** in a mixture with the *N*-Boc ketone **30**. The presence of the ketone can be explained by two degradation reactions. Compound **31** could have underwent either the previously described decarbonylation-decarboxylation combination after ester deprotection (**Scheme 11**), or a retro-Claisen condensation presented in **Scheme 13**.<sup>256</sup>



**Scheme 13:** Proposed retro-Claisen condensation mechanism during the ester deprotection reaction. The high reactivity of the second carbonyl resulted in a nucleophilic attack by hydroxyl group, promoting a C-C bond cleavage and the production of an enolate, subsequently reprotoneated to generate the ketone **30**.

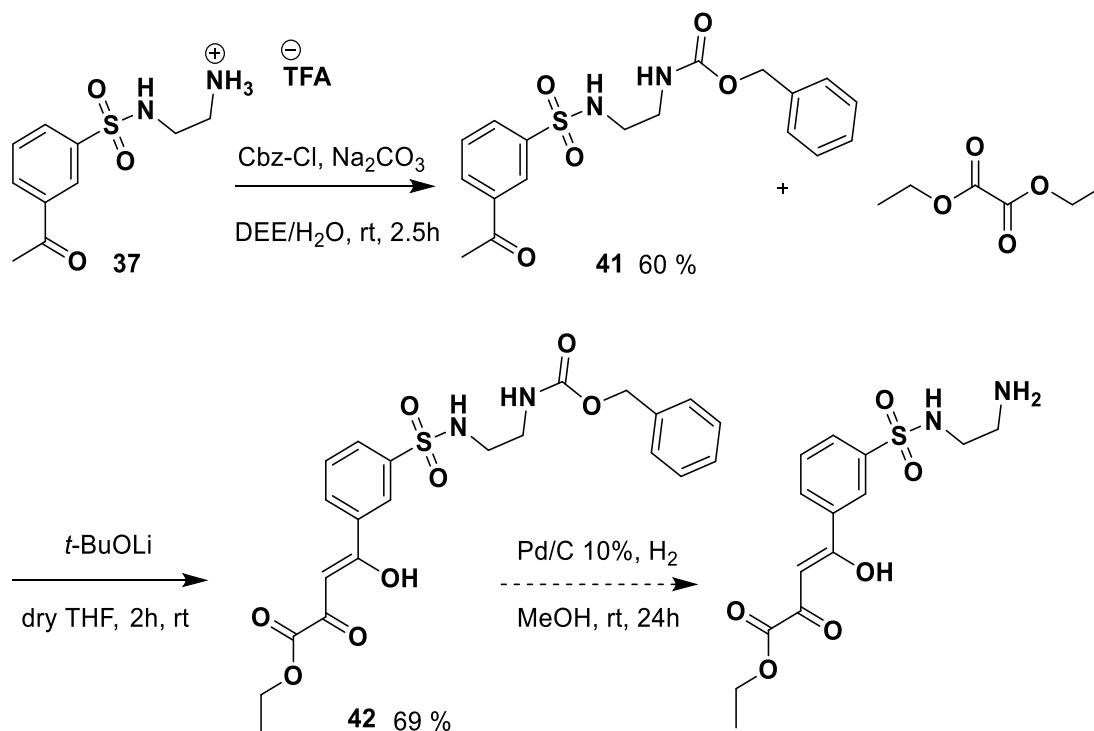
The mixture was further purified by column chromatography and recrystallisation in hexane. The difficulty in the purification and the presence of by-products gave a low yield of 12 % of pure product.  $^1\text{H-NMR}$  spectrum proved the loss of the ethyl protecting group, confirmed later by mass spectrometry (see Appendix 1). This product was later tested for its binding affinity inside ADHs.

#### 4.2.2.3 Use of benzyloxycarbonyl (Cbz) as protecting group of *N*-sulfonyl-ethylenediamine and the deprotection

Given the lack of stability of the  $\beta$ -diketoester functionality under acidic Boc deprotection conditions, a different protective group for the *N*-sulfonyl-ethylenediamine was investigated in parallel. The benzyloxycarbonyl (Cbz) protecting group was chosen due to its stability under basic Claisen condensation condition and its ability to be deprotected without strong acidic conditions, which degrade the  $\beta$ -diketoester. The usual Cbz deprotection conditions imply a catalysed reductive cleavage by hydrogen gas catalysed by palladium on carbon (Pd/C). To the best of our knowledge, no Cbz deprotection has been achieved with  $\beta$ -diketoester group in a molecule. The main risk was the hydrogenation of the aryl ketone as reported by Fadnavis and Radhika.<sup>257</sup>

From the deprotected ketone **37**, the protection of the primary amine by benzyl chloroformate, following the protocol from Song and co-workers,<sup>258</sup> led to the pure Cbz-protected ketone **41** with 60 % yield (**Scheme 14**).  $^1\text{H-NMR}$  spectrum validated

the protocol by showing the signals representing the ketone, the typical Cbz aromatic ring and the ethylenediamine moiety.



**Scheme 14:** Synthetic pathway using Cbz protection. The deprotected ketone **37** was first re-protected by Cbz group before the synthesis of the β-diketoester **42**. The final step was the Cbz deprotection catalysed by Pd/C.

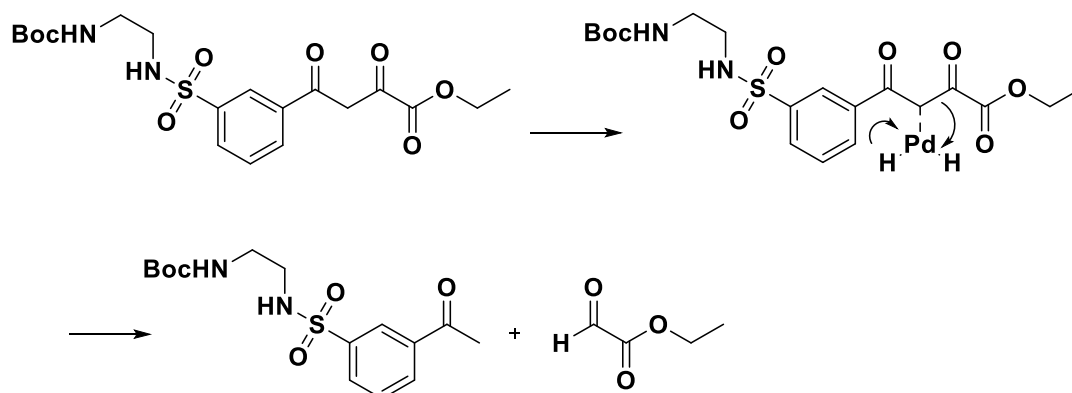
The Claisen condensation of **41** with diethyl oxalate gave the Cbz-β-diketoester **42** product with 69 % yield. This time, the two enols form *cis* and *trans* were not separated by column chromatography. The overall yield was higher than for the previously synthesised Boc-protected compounds (33 % and 37 %), indicating a possible impact of the protecting group geometry on the β-diketoester synthesis step (see Appendix 1 for <sup>1</sup>H-NMR spectra).

Following the protocol from Croft and co-workers,<sup>259</sup> the Cbz protected ketone **41** and the Cbz protected β-diketoester **42** were further subjected to deprotection, using a Pd/C catalyst under hydrogen atmosphere.

The Cbz deprotection of the ketone **41** gave the deprotected ketone **37** with 39 % yield. <sup>1</sup>H-NMR spectrum analysis showed no Cbz characteristic signals and the typical shift of the ethylenediamine protons due to deprotection.

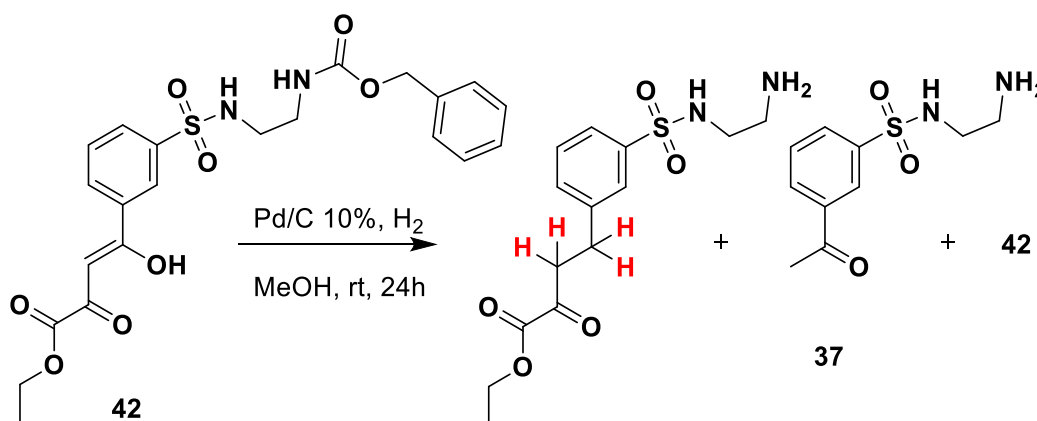
However, with the similar protocol used for the deprotection of ketone **41**, attempts to deprotect the Cbz β-diketoester **42** resulted again in degradation. <sup>1</sup>H-NMR spectrum showed the starting material along with a mixture of by-products. The

deprotected ketone **37** typical signals were spotted. One degradation hypothesis was first the decarboxylation of the ester induced by Pd/C, followed by the already described decarboxylation-decarbonylation combination (**Scheme 11**). Another possible explanation could be the Pd/C catalysed selective C-C bond scission of the diketone by H<sub>2</sub> reaction, resulting in the ketone **37** (**Scheme 15**).<sup>260,261,262</sup>



**Scheme 15:** Proposed mechanism for the hydrogenolysis by H<sub>2</sub> of the  $\beta$ -diketoester anchoring part, catalysed by Pd/C.<sup>263</sup>

Two triplets counting for four hydrogens (2.58 ppm and 2.84 ppm) were also identified in the spectrum at 45 wt% (**Scheme 16**). These signals corresponded to the hydrogenated aryl ketone to hydrocarbon, consistent with the literature on palladium preferential hydrogenolysis of aryl ketones.<sup>264,265,266,257</sup>



**Scheme 16:** Example of observed by-products obtained from the Cbz deprotection of  $\beta$ -diketoester **42** using Pd/C - H<sub>2</sub> hydrogenation.

Only 6 % of the starting material was recovered.

#### 4.2.2.4 Reduction of the $\beta$ -diketoester group.

Given the instability of the substituted  $\beta$ -diketoester, another strategy was the reduction of this entity prior to the amine deprotection. Accordingly, the reduction of the ketone functionalities to hydroxyls was selected as it should prevent the



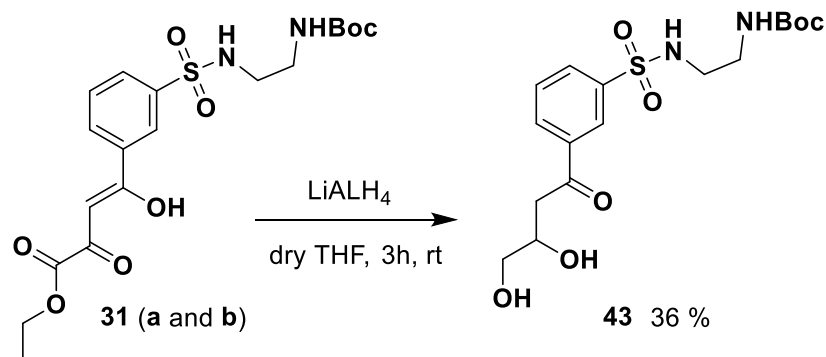
decarbonylation-decarboxylation degradation and it would give access to the second anchor option: product **14**, composed of a diol and a carboxylic acid (**Scheme 2**, page 116). Besides, polyols showed stability under TFA treatment.<sup>267,268</sup>

- **Reduction of the *N*-Boc- $\beta$ -diketoester **31** and the subsequent Boc deprotection**

The prior reduction of the ketones, while keeping the ester intact, will be followed by ester deprotection to a carboxylic acid, to afford the final product **14**. The esters are stabilised by the presence of the second oxygen atom with a lone pair of electrons, avoiding their reduction by weak reagents. The use of a weak reducing agent like sodium borohydride ( $\text{NaBH}_4$ ) was selected, as it promotes the reduction of ketones over esters.<sup>269</sup>

Unfortunately, the treatment with  $\text{NaBH}_4$  of the  $\beta$ -diketoester **31** resulted in a mixture of several products. The  $^1\text{H}$ -NMR and COSY spectra showed a high proportion of the starting material (68 %) and a mixture of by-products was observed, among which a small portion of a  $\beta$ -diketoester with only the aliphatic ketone reduced.

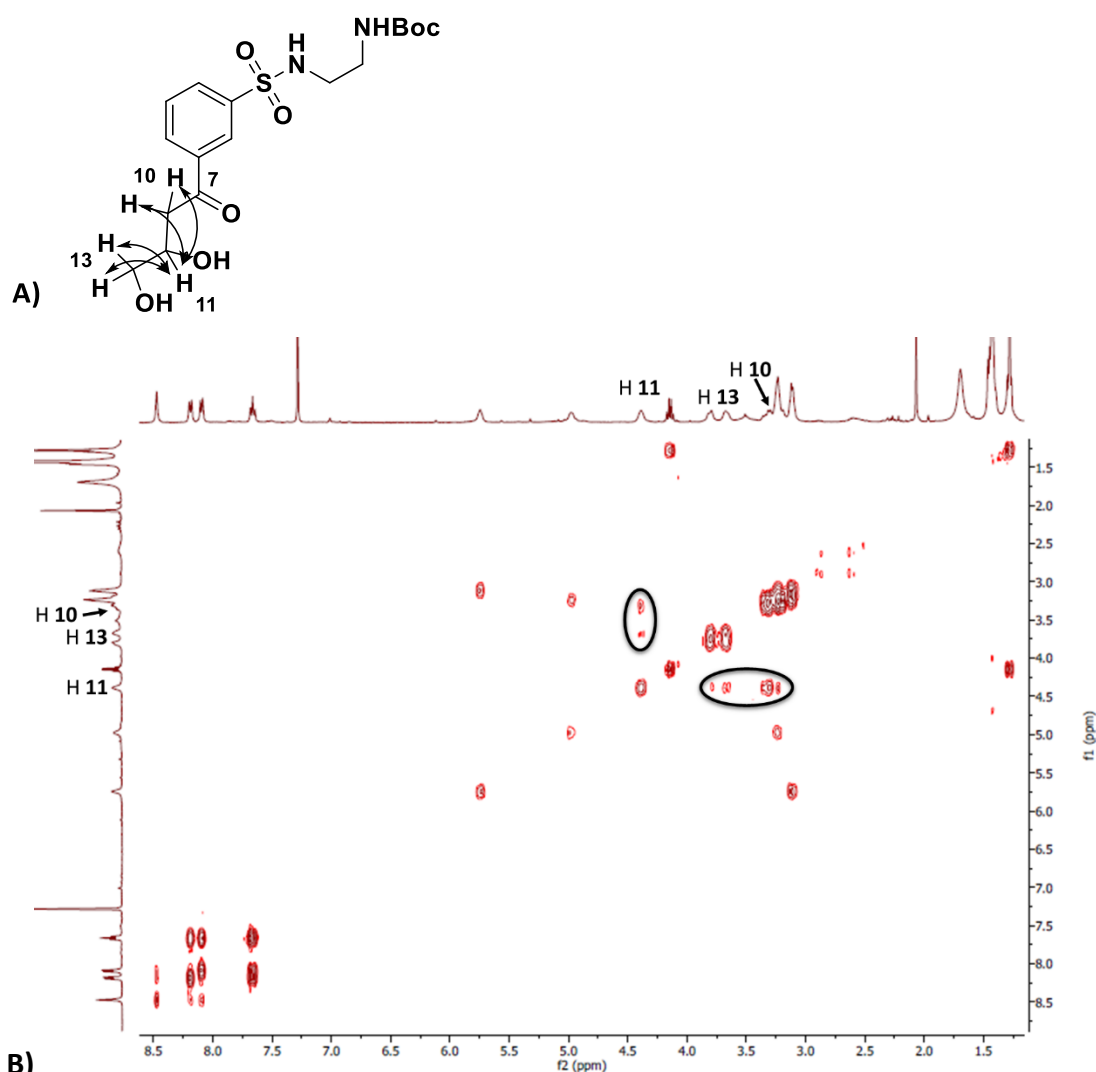
Thus, the second option chosen was to use a stronger reducing agent. Following the optimal conditions according to Ragavan and co-workers,<sup>270</sup> lithium aluminium hydride ( $\text{LiAlH}_4$ , LAH) was tested under inert atmosphere (**Scheme 17**). Product **43** was obtained with a 36 % yield.



**Scheme 17:** The LAH-mediated reduction of compound **31** afforded the Boc-protected keto-diol **43**. Both aliphatic ketone and final ester were reduced to hydroxyl groups.

COSY-NMR and HSQC-NMR experiments confirmed the reduction of both ester and ketone (**Figure 68** and Appendix 1). The ethyl ester signals were not present anymore and two new signals were detected at 3.67 ppm and 3.80 ppm (protons linked to  $\text{C}_{13}$ , where **13** is the carbon number described in **Figure 68**). These protons suggested the

deprotection and reduction of ester into a hydroxyl. In COSY spectrum, these two hydrogen signals correlated to one new hydrogen at 4.87 ppm, representing the proton linked to C<sub>11</sub>. Then, another two hydrogens signal at 3.34 ppm and 3.19 ppm, corresponding to the two vinyl hydrogens hold by C<sub>10</sub>, coupled with the proton on C<sub>11</sub>. No hydrogen signal was displayed at the position C<sub>7</sub>, highlighting the presence of a carbonyl group.

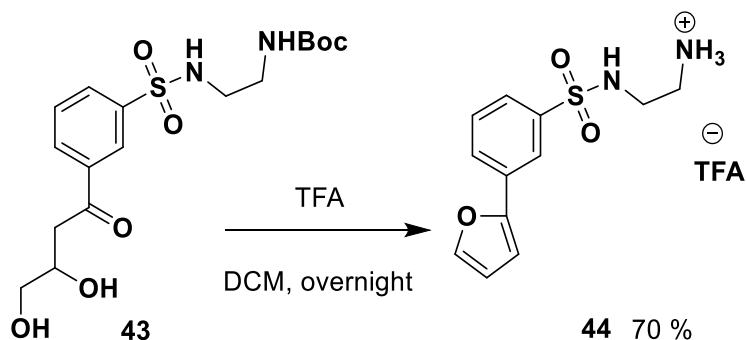


**Figure 68:** A) H-H coupling in product **43**. The arrows represent the couplings made by the H11. B) The COSY-NMR spectrum of product **43** showed the couplings between H11 and H10 and between H11 and H13, highlighted by circles.

As stated by Ragavan and co-workers,<sup>270,271</sup> the ester and the alkyl ketone were chemoselectively reduced to a hydroxyl, while the aryl ketone stayed intact. Their proposed justification to ensure the stability of the aryl ketone was based on the conjugation with the aromatic ring. This lower ketone reactivity avoided any

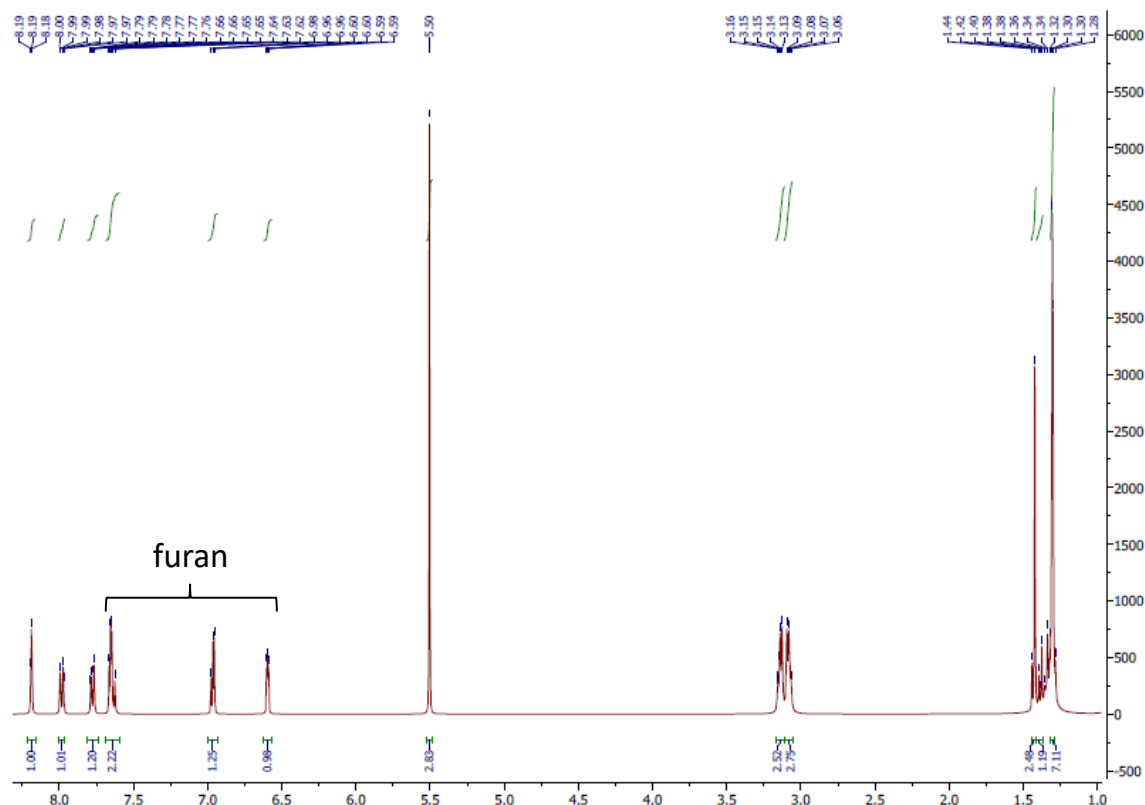
reduction, even with strong agents like LAH. Meanwhile, the aliphatic carbonyl functionalities underwent a reduction to form a diol.

Following the successful reduction of the  $\beta$ -diketoester **31** to the ketodiol **43**, the TFA-mediated *N*-Boc deprotection was again attempted, using the previously described protocol.



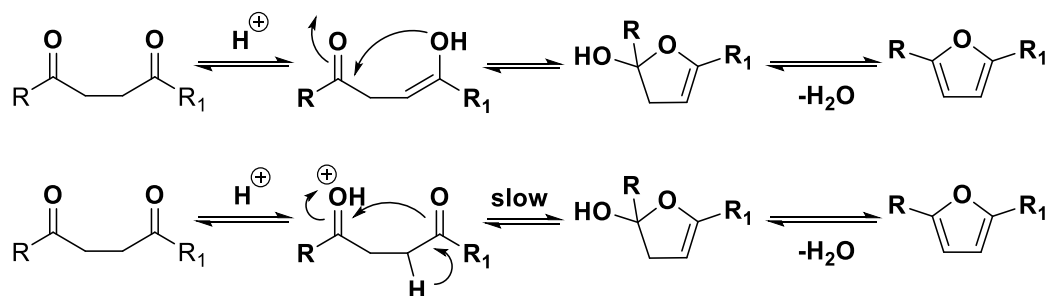
**Scheme 18:** TFA-mediated *N*-Boc deprotection of the ketodiol **43**, using the previously described protocol. The resulting product **44** was a deprotected furan.

This time, no degraded by-products or starting material were observed by TLC or  $^1\text{H}$ -NMR. However, the expected final product **14** was not detected. The analysis of  $^1\text{H}$ -NMR, COSY-NMR, and  $^{13}\text{C}$ -NMR spectra showed the presence of three new aromatic signals coupling together (between 6.59 ppm and 7.77 ppm, **Figure 69** and Appendix 1). Besides, hydrogen signals composing the ketodiol part of the starting material **43** could not be found ( $\text{C}_{10}$ ,  $\text{C}_{11}$  and  $\text{C}_{13}$ , **Figure 68**). The NMR spectra, supported by mass spectrometry, allowed to conclude on a cyclisation of the diol and ketone, due to the use of acidic conditions. The TFA-mediated cyclisation afforded the furan **44** with a 70 % yield.



**Figure 69:**  $^1\text{H}$ -NMR spectrum of product **44**. The three new observed aromatic signals correspond to furan hydrogens.

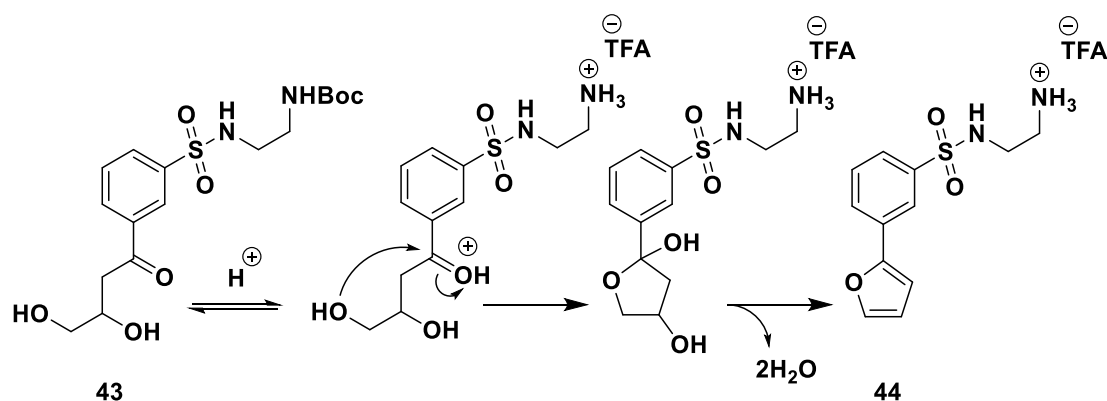
A proposed explanation of the furan synthesis would be the occurrence of the acid-catalysed Paal-Knorr reaction from 1,4-diketones, discovered in 1884 (**Scheme 19**).<sup>272,273</sup> Even though the reaction is known for more than a century, few studies have explored the reaction mechanism, and the pathway is still debated. A commonly accepted mechanism is the ring-closure of a rapidly formed mono-enol. However, little experimental work has been carried out to validate this pathway based on hypothesis from pyrrole synthesis experiments.<sup>274,275</sup> Based on experimental kinetic studies on furan, Amarnath and Amarnath later proposed a new mechanism involving a concerted cyclisation by attack of a protonated carbonyl group by the readily formed enol from the other carbonyl (**Scheme 19**).<sup>276</sup> This is then followed by an irreversible dehydrative cyclisation to form the final furan.



**Scheme 19:** Top: generally accepted Paal-Knorr mechanism for the acidic catalysed cyclisation of diketone into furan. Bottom: Reported Paal-Knorr pathway by Amarnath and Amarnath.<sup>276,277</sup>

The synthesis of furan derivatives from alcohols was also demonstrated by Weedon and co-workers in acidic conditions, due to cyclisation of a 1,4-keto-alcohol intermediate, with no mechanism determination.<sup>278,279,280</sup> Furthermore, the specific TFA-catalysed cyclisation of a *tert*-butyl acetoacetate was proven by Stauffer and Neier.<sup>281</sup>

Based on literature results and mechanisms hypotheses, a proposed pathway for the TFA-catalysed cyclisation of product **43** is presented in **Scheme 20**.



**Scheme 20:** Proposed mechanism of the TFA-catalysed furan cyclisation to form the furan product **44**. The mechanism was based on reported Paal-Knorr pathways.<sup>276,277</sup>

Based on Amarnath and Amarnath suggestions, the acidic condition led to the protonation of the aryl carbonyl, inducing the electrophilic attack of the 4-hydroxyl on the protonated carbonyl centre. The cyclisation was followed by irreversible dehydration due to aromatisation.

The newly synthesised product **44** represented the perspective of a new catalyst complex and was used in the next step of metalation.

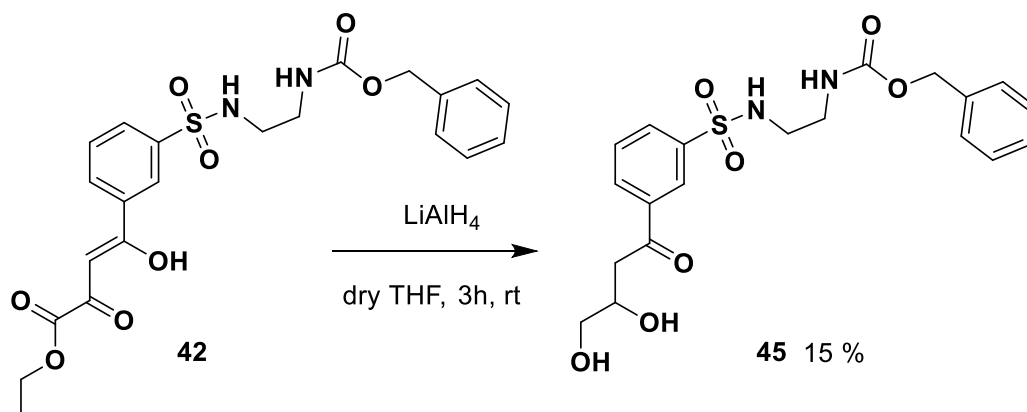
In the previous TFA-mediated Boc deprotection attempts, the  $\beta$ -diketoester products **31** and **40** and the ketoacid **38** did not undergo cyclisation.

First, they were all composed of a 1,3-diketone group and the Paal-Knorr reaction needs a 1,4-dicarbonyl reactant to cyclise. Thus, in the  $\beta$ -diketoester and the ketoacids groups, only the carbonyl from the ester and the carboxylic acid would have reacted with the aryl ketone.

To the best of our knowledge, no evidence of an ester carbonyl involved in Paal-Knorr reaction was found in literature. Deprotection is necessary to obtain the reactive carboxylic acid. Stauffer and Neier showed an example of a furan cyclisation between ketone and carboxylic acid, using a TFA-catalysed reaction.<sup>281</sup> Their tert-butyl acetoacetate products were deprotected by TFA to give the corresponding carboxylic acid, which attack the keto function, leading to a cyclisation. However, they confirmed the lack of reactivity when treating similar product as ketoacid **38** with TFA. The low electrophilic activity of the aryl ketone carbonyl is not enough for a possible attack of the carboxylic acid.

- **Reduction of the Cbz- $\beta$ -diketoester **42** and the subsequent Cbz deprotection**

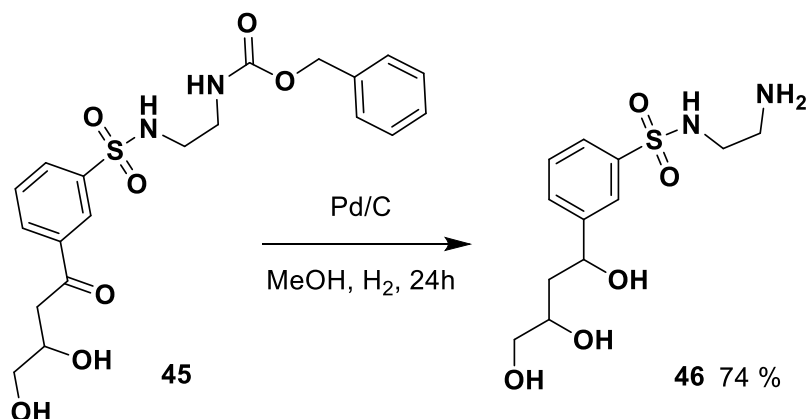
To circumvent the cyclisation induced by acidic conditions, the Pd/C - H<sub>2</sub> conditions were again tested for Cbz deprotection of the reduced  $\beta$ -diketoester. The Cbz protected  $\beta$ -diketoester **42** was first reduced by LAH, following the previously used protocol (**Scheme 21**). This time, after column chromatography purification, the Cbz-ketodiol **45** was obtained with 15 % yield.



**Scheme 21:** The LAH-mediated reduction of the Cbz-protected product **42** afforded the Cbz-protected keto-diol **45**. Both aliphatic ketone and final ester were reduced to hydroxyl groups.

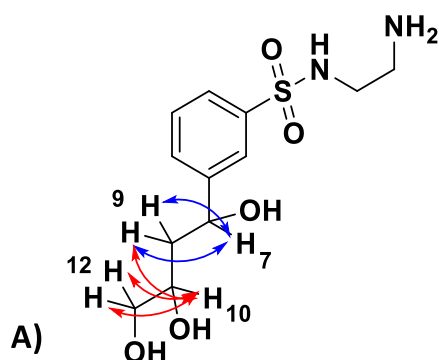
The yield was influenced by the purification (recovery of impure product). For further optimisation of the reaction, a solution could be the use of a weaker reducing agents (NaBH<sub>4</sub>), or lower LAH concentrations (literature references advised 3 eq instead of the protocol 4 eq).<sup>282,283</sup>

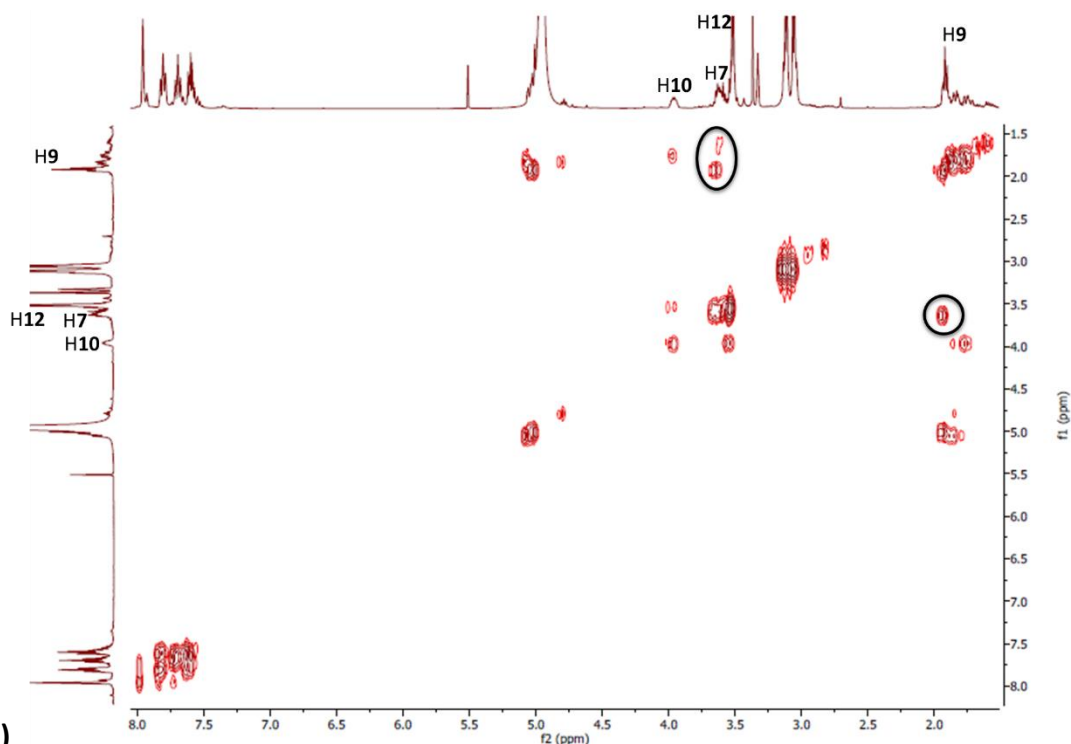
The removal of the Cbz group was then performed using Pd/C - H<sub>2</sub>, applying the same condition as previously described:



**Scheme 22:** Synthesis of the triol **46** by Cbz deprotection of product **45**, using a reductive cleavage by hydrogen gas catalysed by palladium on carbon.

The reductive cleavage of the Cbz group was successful, without Paal-Knorr cyclisation. The typical aromatic ring signals were not observed on the <sup>1</sup>H-NMR spectrum (Appendix 1). Under these deprotection conditions, the reduction of the aromatic ketone was confirmed by NMR. A new hydrogen signal emerged in the <sup>1</sup>H-NMR spectrum, corresponding to the proton linked to C<sub>7</sub>, represented in **Figure 70**. A COSY-NMR spectrum confirmed the coupling between both C<sub>7</sub> and C<sub>9</sub> protons (circled, **Figure 70**). In addition, the absence of carbonyl signals was spotted on the <sup>13</sup>C-NMR spectrum.



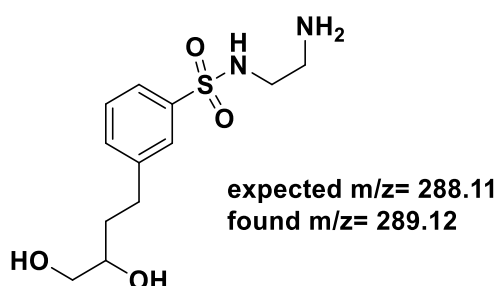


**B)**

**Figure 70:** A) H-H coupling in product **46**. The blue arrows represent the couplings made by the H7. The red arrows represent the couplings made by the H10. B) The COSY-NMR spectrum of product **46** showed the coupling of the new H7 with the two neighbours H9, highlighted by a circle.

These data confirmed the catalytic hydrogenation of the aromatic ketone using Pd/C - H<sub>2</sub> conditions, already described in literature.<sup>266,284</sup> The final triol **46** was obtained with 74 % yield and taken to the next step.

A 17 % yield of by-product was also found and characterised by ESI-MS and <sup>1</sup>H-NMR. The product corresponds to the total reduction of the aryl ketone by hydrogenation from Pd/C - H<sub>2</sub> conditions:



**Figure 71:** By-product of the total reduction of **45** aromatic ketone, under Cbz deprotection conditions using Pd/C - H<sub>2</sub>.

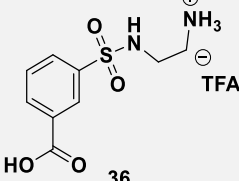
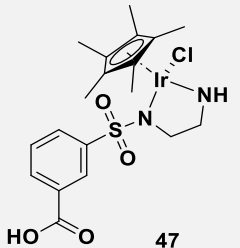
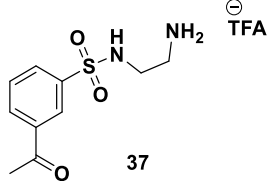
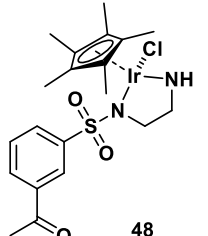
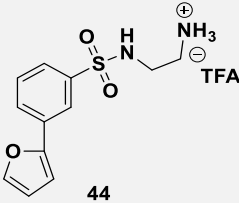
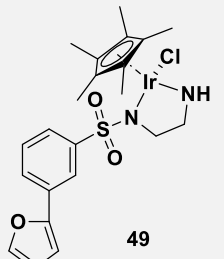
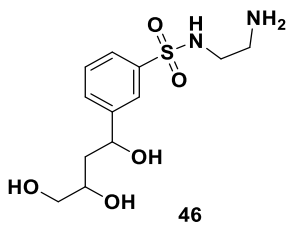
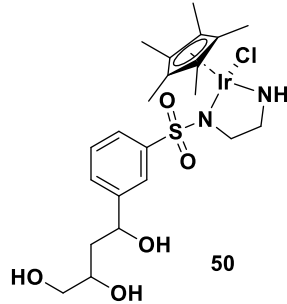
#### 4.2.3 Transition metal complexation of all deprotected compounds

All deprotected N-sulfonyl-ethylenediamine products (**36**, **37**, **44** and **46**) were subjected to iridium metalation to form the final piano-stool complexes. The protocol was based on the modified work of Letondor and co-workers.<sup>58</sup> The



pentamethylcyclopentadienyl iridium dichloride dimer ( $[\text{IrCl}_2\text{Cp}^*]_2$ ) was first suspended in methanol while the deprotected compounds, dissolved in methanol, were added to the suspension. A diphase water/methanol mixture was used to dissolve compound **36**. All reactions were performed under nitrogen at room temperature until completion and the mixture was purified by precipitation in hexane. The different transition metal catalysts obtained were:

**Table 12:** Metalation of the deprotected compounds **36**, **37**, **44** and **46** using  $[\text{Ir}_2\text{Cl}_2\text{Cp}^*]_2$  and following the modified protocol from Letondor and co-workers.<sup>58</sup>

Substrate	Conditions	Product	Yield
 <b>36</b>	MeOH/water, 1h, rt	 <b>47</b>	55 %
 <b>37</b>	MeOH, rt	 <b>48</b>	20 %
 <b>44</b>	MeOH, 1h, rt	 <b>49</b>	43 %
 <b>46</b>	MeOH, 1h, rt	 <b>50</b>	43 %

In all  $^1\text{H}$ -NMR spectra, the metalation was confirmed by a shift of the two  $\text{CH}_2$  hydrogens of the *N*-sulfonyl-ethylenediamine, from approximately 3.15-3.05 ppm to

2.60-2.50 ppm. Upon binding of the metal, the Cp\* signal at around 1.76 ppm was also identified (see Appendix 1).

Affinity for the ADHs scaffold was then tested to compare hydrophilic (**50**), hydrophobic (**49**) and short chain (**47**) catalysts potential.

### 4.3 Conclusion

Following the computational design, a divergent synthetic strategy was designed to produce the selected hit compounds. The strategy was optimised to represent an efficient pathway with limited steps. Attempts to obtain the desired anchoring part were unsuccessful at the final step of *N*-Boc deprotection. Several attempts with changes in the reaction conditions, the anchoring part substitutions, or the use of an alternative protecting group were all non-successful. A degradation of the  $\beta$ -diketoester always occurred before the necessary deprotecting group cleavage.

A final strategy was based on reducing the  $\beta$ -diketoester, prior to deprotection. This strategy succeeded, with the synthesis of two new anchoring products: a furan **44** and a triol **46**. The product **46** was close enough to the predicted structure **14** (only a carboxylic acid replaced by hydroxyl) to be considered as an anchor for the Ir metal. Thus, both anchors and an intermediate (product **36**) were metalated with Ir to afford the final characterised transition metal complexes: [IrClCp\*]-benzoic acid **47**; [IrClCp\*]-furan **49**; [IrClCp\*]-triol **50**.

These new catalysts differed from the expected products, but still hold the functional groups and overall structures likely to bind inside the enzyme.

Three metal catalysts complexes were thus synthesised, as well as their three anchoring ligands. Another non Boc-deprotected ligand, product **38**, was also synthesised to test the binding affinity of ligand similar to computationally selected lead structure **14**.

To create an AM, all the different metal complexes were first to be tested for their supramolecular binding with HLADH-WT, TbADH-WT and TbADH-5M. As new synthesised catalysts, their catalytic activity in the imine reduction was also subsequently investigated.

## 4.4 Experimental methods

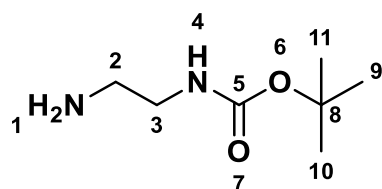
### 4.4.1 Materials

All chemical reagents were purchased as analytical grade from Sigma Aldrich and Fisher Scientific. The commercially available chemicals were used without purification unless otherwise stated. All aqueous solutions were prepared using deionised water.

Analytical thin layer chromatography (TLC) was carried out on aluminium backed plates coated with Merck Kieselgel 60 GF254 and visualised under UV light at 254 and/or 360 nm. Chemical staining was also routinely used, with aqueous basic potassium permanganate solution. Flash chromatography was carried out using Davisil silica 60 Å, with eluent specified. Infrared spectra were recorded using Bruker FTIR spectrometer ALPHA II over the range 4000-600 cm<sup>-1</sup>. NMR spectra were run at 298 K using a Bruker AV(III)400, AV400. DPX400 (400 MHz <sup>1</sup>H frequency, 100 MHz <sup>13</sup>C frequency). Chemical shifts are quoted in parts per million (ppm), referenced to the residual deuterated solvent quoted as internal standard. Coupling constant *J* are in Hz. Multiplicity of the signals is abbreviated as follow: s, singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublet; dt, doublet of triplet; m, multiplet; br, broad. In the <sup>13</sup>C spectra, signals corresponding to C, CH, CH<sub>2</sub> and CH<sub>3</sub> were assigned from DEPT experiments. Mass spectra were recorded using a Bruker MicroTOF 61 mass spectrometer using electrospray ionisation (ESI).

### 4.4.2 Methods

#### **tert-butyl (2-aminoethyl)carbamate (26)**

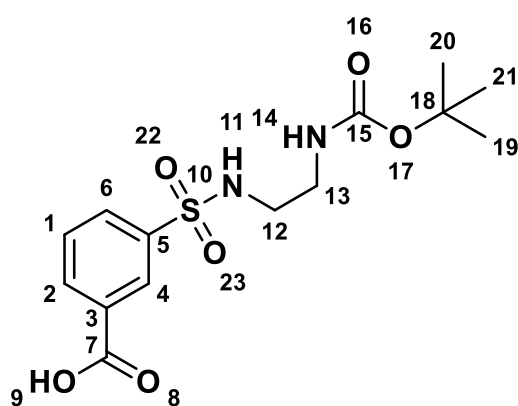


To a cooled solution of commercial ethylenediamine (9.2 eq, 171.5 mmol, 11 mL) in chloroform (200 mL) was added dropwise under stirring a commercial di-*tert*-butyl dicarbonate (1 eq, 18.6 mmol, 4.2 mL). The solution was then stirred overnight at room temperature.

The solution was washed successively with brine, 1 M NaOH three times, brine and deionized water. The organic layers were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and the solvent evaporated under *vacuum* to obtain the known product **26** as a yellow precipitate (2.48 g, 83 %).<sup>285</sup> The characterisation data were consistent with those published in the literature.<sup>286</sup>

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 5.27 (br.s, 1H, H<sub>4</sub> NH), 3.12 (dt, 2H, H<sub>3</sub> CH<sub>2</sub>), 2.02 (t, 2H, H<sub>2</sub> CH<sub>2</sub>), 2.03 (s, 2H, NH<sub>2</sub>), 1.40 (s, 9H, H<sub>9</sub>-H<sub>10</sub>-H<sub>11</sub> CH<sub>3</sub>). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 28.3 (3C); 41.7; 43.1; 79.0; 156.3.

### **3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfamoyl)benzoic acid (28)**

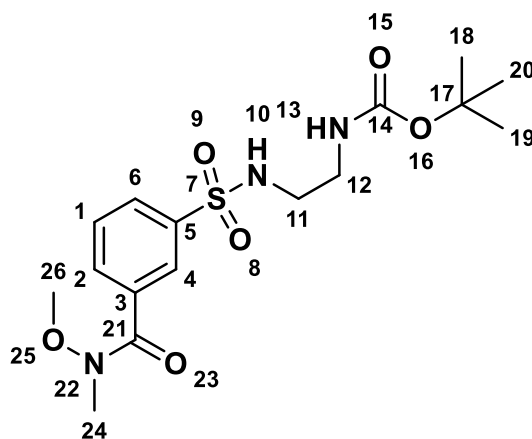


This compound was synthesised by adapting a previously described protocol.<sup>220</sup> To a cooled solution of the product **26** (1.2 eq, 9.3 mmol, 1.5 g) in dichloromethane (36 mL) was added dropwise under stirring trimethylamine (1.5 eq, 14 mmol, 2 mL) and a solution of the commercially available chlorosulfonyl benzoic acid 3-(chlorosulfonyl)benzoic acid (1 eq, 7.71 mmol, 1.7 g) in dichloromethane (48 mL). The solution was stirred overnight at room temperature and monitored by TLC (CH<sub>2</sub>Cl<sub>2</sub> 95/5 CH<sub>3</sub>OH + traces of acetic acid). The solution was evaporated under *vacuum* to obtain the known product **28** as a white powder (2.54 g, 92 %).

<sup>1</sup>H NMR (400 MHz, Acetone): δ = 8.50 (s, 1H, H<sub>4</sub> aryl), 8.30 (d, 1H, J=8.0 Hz, H<sub>2</sub> aryl), 8.28 (d, 1H, J=8.0 Hz, H<sub>6</sub> aryl), 8.11 (t, 1H, J=8.0 Hz, H<sub>1</sub> aryl), 6.80 (m, 1H, H<sub>11</sub> NH); 6.07 (m, 1H, H<sub>14</sub> NH); 3.19 (m, 2H, H<sub>12</sub> CH<sub>2</sub>); 3.05 (m, 2H, H<sub>13</sub> CH<sub>2</sub>); 1.39 (s, 9H, H<sub>19</sub>-H<sub>20</sub>-H<sub>21</sub> CH<sub>3</sub>). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 28.33 (3C); 40.28; 43.83; 61.34; 126.92; 128.89; 129.05; 132.35; 135.07; 140.06; 168.00 (2C).

ESI-MS positive mode calculated for  $C_{14}H_{20}O_6N_2S_1$  344.1042, found 367.0934 ( $[M+Na^+]$ ).

**tert-butyl(2-((3(methoxy(methyl)carbamoyl)phenyl)sulfonamido)ethyl)carbamate**  
**(29)**



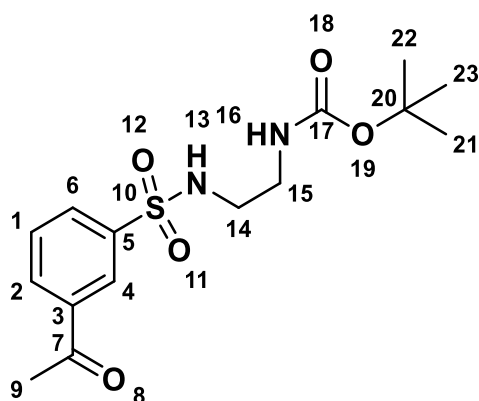
This compound was synthesised by adapting a previously described protocol.<sup>225</sup> To a cooled solution of the acid **28** (1 eq, 0.145 mmol, 50 mg) in anhydrous dichloromethane (3 mL) was added dropwise under stirring thionyl chloride (7 eq, 1.016 mmol, 73  $\mu$ L). The solution was then stirred at room temperature for 3h. The reaction was cooled at 0°C and N,N-diisopropylethylamine (4 eq, 0.58 mmol, 0.101 mL) and N,O dimethylhydroxylamine (2 eq, 0.290 mmol, 28 mg) were added dropwise. The reaction was then stirred at room temperature overnight. The reaction was treated with saturated aqueous  $NaHCO_3$  (pH=7) then washed twice with water and dried over  $Na_2SO_4$ , filtered, and evaporated under *vacuum*.

The crude extract was purified by flash chromatography column on silica gel (ethyl acetate 60/40 hexane) to afford product **29** (20 mg, 35 %).

$^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta$  = 8.18 (s, 1H,  $H_4$  aryl), 7.95-7.93 (d, 1H,  $J$ =8.0 Hz,  $H_2$  aryl), 7.90-7.88 (d, 1H,  $J$ =8.0 Hz,  $H_6$  aryl), 7.58-7.54 (t, 1H,  $J$ =8.0 Hz,  $H_1$  aryl), 5.67 (m, 1H,  $H_{10}$  NH); 4.99 (m, 1H,  $H_{13}$  NH); 3.54 (s, 3H,  $H_{26}$   $CH_3$ ); 3.38 (s, 3H,  $H_{24}$   $CH_3$ ); 3.23-3.19 (dd, 2H,  $J_1$ =4.0 Hz,  $J_2$ =12 Hz,  $H_{11}$   $CH_2$ ); 3.08-3.03 (dd, 2H,  $J_1$ =4.0 Hz,  $J_2$ =12 Hz,  $H_{12}$   $CH_2$ ); 1.41 (s, 9H,  $H_{18}$ - $H_{19}$ - $H_{20}$   $CH_3$ ).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ ):  $\delta$  = 28.33 (3C); 31.93; 40.28; 43.80; 61.34; 79.94; 126.92; 128.89; 129.05; 132.35; 135.07; 140.06; 168.00 (2C).

ESI-MS positive mode calculated for  $C_{16}H_{25}O_6N_3S_1$  387.1464, found 388.1545 ( $[M+H^+]$ ), 410.1348 ( $[M+Na^+]$ ) and 797.2839 ( $[2M+Na^+]$ ).

**tert-butyl (2-((3-acetylphenyl)sulfonamido)ethyl)carbamate (30)**

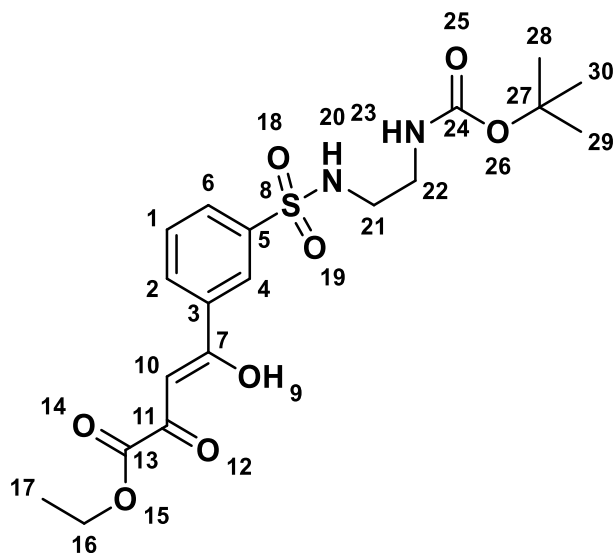


This compound was synthesised by adapting a previously described protocol.<sup>226</sup> To a cooled solution of *tert*-butyl (2-((3-(methoxy(methyl)carbamoyl)phenyl)sulfonamido)ethyl)carbamate **29** (1 eq, 0.098 mmol, 38 mg) in anhydrous THF (2 mL), under nitrogen, was added dropwise under stirring a solution of methylmagnesium bromide (3 eq, 0.294 mmol, 0.15 mL). The solution was then stirred at room temperature for 3h. The reaction was monitored by TLC (ethyl acetate 60/40 hexane). The solution was treated with saturated aqueous NaHCO<sub>3</sub> (pH=7) and extracted twice with ethyl acetate. The organic layers were washed with brine and water then dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and evaporated under *vacuum*. The crude extract was purified by flash chromatography column on silica gel (ethyl acetate 60/40 hexane) to afford product **30** as a yellow oil (20 mg, estimated 54 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 8.41 (s, 1H, H<sub>4</sub> aryl), 8.15-8.13 (d, 1H, *J*=8.0 Hz, H<sub>2</sub> aryl), 8.06-8.04 (d, 1H, *J*=8.0 Hz, H<sub>6</sub> aryl), 7.63 (t, 1H, *J*=8.0 Hz, H<sub>1</sub> aryl), 6.04 (m, 1H, H<sub>13</sub> NH); 5.13 (m, 1H, H<sub>16</sub> NH); 3.22 (dd, 2H, *J*<sub>1</sub>=4.0 Hz, *J*<sub>2</sub>=12 Hz, H<sub>14</sub> CH<sub>2</sub>); 3.07 (dd, 2H, *J*<sub>1</sub>=4.0 Hz, *J*<sub>2</sub>=12 Hz, H<sub>15</sub> CH<sub>2</sub>); 2.65 (s, 3H, H<sub>9</sub> CH<sub>3</sub>); 1.39 (s, 9H, H<sub>21</sub>-H<sub>22</sub>-H<sub>23</sub> CH<sub>3</sub>). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  = 28.33 (3C); 26.70; 40.28; 43.83; 61.34; 126.92; 128.89; 129.05; 132.35; 135.07; 140.06; 168.00 (2C).

ESI-MS positive mode calculated for C<sub>15</sub>H<sub>22</sub>O<sub>5</sub>N<sub>2</sub>S<sub>1</sub> 342.1249, found 365.1144 ([M+Na<sup>+</sup>]), 685.2577 ([2M+H<sup>+</sup>]) and 707.2407 ([2M+Na<sup>+</sup>]).

**Ethyl 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate (31)**



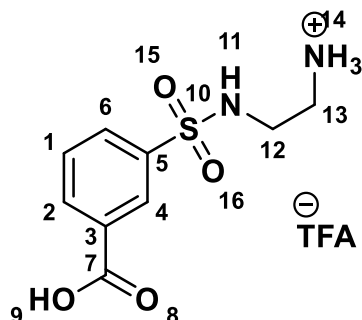
This compound was synthesised by adapting a previously described protocol.<sup>233</sup> To a cooled solution of *t*-BuOLi 1 M (2.5 eq, 1.65 mmol, 1.65 mL) in anhydrous THF (5 mL) was added dropwise under nitrogen diethyl oxalate (2 eq, 1.36 mmol, 0.2 mL). The solution was stirred at 0 °C for 10 minutes before the ketone **30** (1 eq, 0.66 mmol, 226 mg) in anhydrous THF (4 mL) was added dropwise. The mixture was stirred for 1.5h at room temperature. The reaction was monitored by TLC (ethyl acetate 50/50 hexane) and treated with saturated aqueous NaHCO<sub>3</sub> (pH=6). The aqueous phase was extracted twice with dichloromethane. The organic layers were washed with brine and water, dried over Na<sub>2</sub>SO<sub>4</sub> and filtered. The solvent was removed *in vacuum* and the resulting residue was purified by flash chromatography column on silica gel (DCM 60/40 MeOH). The title compound **31** was obtained as a yellow oil (96 mg, estimated 26 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 8.62 (s, 1H, H<sub>4</sub> aryl); 8.04 (d, 1H, *J*=8.0 Hz, H<sub>2</sub> aryl); 7.93 (d, 1H, *J*=8.0 Hz, H<sub>6</sub> aryl); 7.52 (t, 1H, *J*=8.0 Hz, H<sub>1</sub> aryl); 6.73 (s, 1H, H<sub>10</sub> CH); 5.68 (m, 1H, H<sub>20</sub> NH); 5.24 (m, 1H, H<sub>23</sub> NH); 4.34 (dd, 2H, *J*<sub>1</sub>=8.0 Hz, *J*<sub>2</sub>=16 Hz, H<sub>16</sub> CH<sub>2</sub>); 2.22 (m, 2H, H<sub>21</sub> CH<sub>2</sub>); 2.07 (m, 2H, H<sub>22</sub> CH<sub>2</sub>); 1.43 (s, 9H, H<sub>28</sub>-H<sub>29</sub>-H<sub>30</sub> CH<sub>3</sub>); 1.35 (s, 3H, H<sub>17</sub> CH<sub>3</sub>).

<sup>13</sup>C (100 MHz, CDCl<sub>3</sub>): δ= 14.11, 28.30 (3C); 40.30; 44.14; 59.02; 62.91; 97.92; 126.24; 129.96; 131.53; 131.57; 135.97; 141.32; 161.79; 171.01; 188.42.

ESI-MS positive mode calculated for  $C_{19}H_{26}O_8N_2S_1$  442.1410, found 465.1298 ( $[M+Na^+]$ ), 885.2881 ( $[2M+H^+]$ ) and 907.2683 ( $[2M+Na^+]$ ).

**(4-carboxyphenylsulfonyl)-ethylenediamine acid (36)**

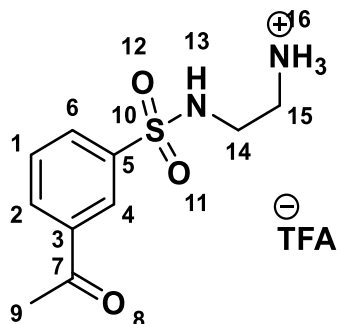


To a cooled solution of the acid **28** (1 eq, 2.87 mmol, 987 mg) in dichloromethane (47 mL) was added dropwise under stirring trifluoroacetic acid (12.7 eq, 36.45 mmol, 2.8 mL). The solution was then stirred overnight at room temperature and monitored by TLC ( $CH_2Cl_2$  95/5  $CH_3OH$  + traces of acetic acid). The solution was evaporated under *vacuum* and co-distilled three times with dichloromethane to obtain the known product **36** as a white powder (358 mg, estimated 30 %).<sup>250</sup>

$^1H$  NMR (400 MHz,  $DMSO-d_6$ ):  $\delta$  = 8.35 (s, 1H,  $H_4$  aryl); 8.23-8.21 (d, 1H,  $J=8.0$  Hz,  $H_2$  aryl), 8.10 (m, 1H,  $H_{11}$  NH); 8.06-8.04 (d, 1H,  $J=8.0$  Hz,  $H_6$  aryl); 7.87 (m, 3H,  $H_{14}$   $NH_3$ ); 7.80-7.76 (t, 1H,  $J=8.0$  Hz,  $H_1$  aryl); 2.98-2.94 (m, 2H,  $H_{12}$   $CH_2$ ); 2.89-2.86 (m, 2H,  $H_{13}$   $CH_2$ ).  $^{13}C$  NMR (100 MHz,  $DMSO-d_6$ ):  $\delta$  = 38.87; 40.33; 127.74; 130.58; 131.17; 132.45; 133.77; 140.55; 210.54.

ESI-MS positive mode calculated for  $C_9H_{13}O_4N_2S_1^+$  245.0591, found 245.0589 ( $[M^+]$ ) and 267.0395 ( $[M+Na^+]$ ).

**3-acetyl-N-(2-aminoethyl)benzenesulfonamide (37)**



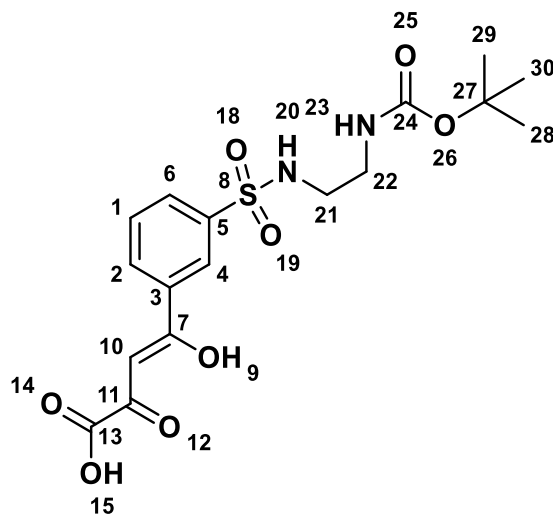


The same protocol as for the synthesis of the product **36** was applied to the ketone **30**.<sup>250</sup> The known compound **37** was collected as a yellow powder (24 mg, estimated 65 %).

<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>): δ = 8.31 (s, 1H, H<sub>4</sub> aryl); 8.27 (dt, 1H, *J*=8.0 Hz, H<sub>2</sub> aryl); 8.13 (t, 1H, *J*=8.0 Hz, H<sub>13</sub> NH); 8.06 (dt, 1H, *J*=8.0 Hz, H<sub>6</sub> aryl); 7.91 (m, 3H, H<sub>16</sub> NH<sub>3</sub>); 7.81 (t, 1H, *J*=8.0 Hz, H<sub>1</sub> aryl); 2.97 (m, 2H, H<sub>14</sub> CH<sub>2</sub>); 2.88 (m, 2H, H<sub>15</sub> CH<sub>2</sub>); 2.66 (s, 3H, H<sub>9</sub> CH<sub>3</sub>). <sup>13</sup>C NMR (100 MHz, DMSO-d<sub>6</sub>): δ = 27.26; 38.94; 40.30; 126.07; 130.59; 131.17; 131.60; 132.58; 139.20; 209.63.

ESI-MS positive mode calculated for C<sub>10</sub>H<sub>15</sub>N<sub>2</sub>O<sub>3</sub>S<sub>1</sub><sup>+</sup> 242.0798, found 243.0796 ([M<sup>+</sup>]) and 265.0611 ([M<sup>+</sup> + Na<sup>+</sup>]).

**4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoic acid (38)**



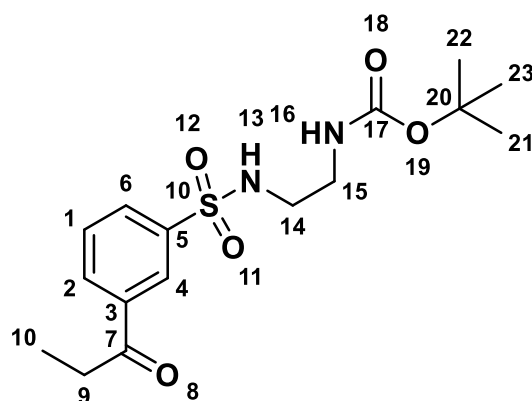
This compound was synthesised by adapting a previously described protocol.<sup>255</sup> To a cooled solution of the *N*-Boc-β-diketoester **31** (1 eq, 0.78 mmol, 343 mg) in a mix H<sub>2</sub>O 3:6 MeOH (20 mL), was added dropwise under stirring NaOH 4 M (263 μL). The solution was then stirred for 30 min at room temperature and monitored by TLC (DCM 90/10 MeOH). The reaction was treated with 6 N H<sub>2</sub>SO<sub>4</sub> (pH=6) and MeOH was evaporated under *vacuum* before the water layer was extracted twice with ethyl acetate. The organic layers were washed with brine and water, dried over Na<sub>2</sub>SO<sub>4</sub> and filtered. The solvent was removed *in vacuum* and the resulting residue was purified by flash chromatography column (acetonitrile 90/10 MeOH) and dissolved in 2 mL of

dichloromethane before recrystallisation in a excess of hexane. The title compound **38** was collected as an orange powder (39 mg, 12 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  = 8.49 (s, 1H,  $\text{H}_4$  aryl); 8.30 (d, 1H,  $J=8.0$  Hz,  $\text{H}_2$  aryl); 8.12 (d, 1H,  $J=8.0$  Hz,  $\text{H}_6$  aryl); 7.78 (t, 1H,  $J=8.0$  Hz,  $\text{H}_1$  aryl); 7.20 (s, 1H,  $\text{H}_{10}$ ); 3.12 (m, 2H,  $\text{H}_{21}$   $\text{CH}_2$ ); 2.97 (m, 2H,  $\text{H}_{22}$   $\text{CH}_2$ ); 1.42 (s, 9H,  $\text{H}_{28}\text{-H}_{29}\text{-H}_{30}$   $\text{CH}_3$ ).

ESI-MS negative mode calculated for  $\text{C}_{17}\text{H}_{22}\text{O}_8\text{N}_2\text{S}_1$  414.1097, found 413.1012 ( $[\text{M}-\text{H}^+]$ ).

**tert-butyl (2-((3-propionylphenyl)sulfonamido)ethyl)carbamate (39)**

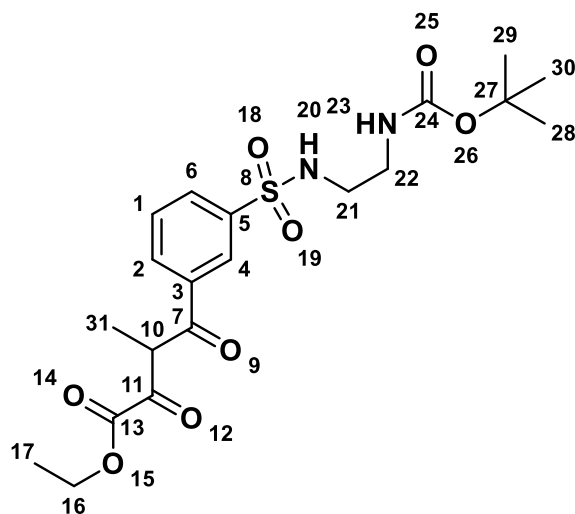


This compound was synthesised by adapting a previously described protocol.<sup>226</sup> To a cooled solution of the *N*-Boc amide **29** (1 eq, 1.55 mmol, 138 mg) in anhydrous THF (53 mL), was added dropwise, under nitrogen, a solution of ethylmagnesium bromide (3 eq, 4.65 mmol, 1.55 mL). The mixture was then stirred at room temperature for 2h30 and monitored by TLC (ethyl acetate 60/40 hexane). The solution was treated with saturated aqueous  $\text{NaHCO}_3$  (pH=7) and extracted twice with ethyl acetate. The organic layers were washed with brine and water then dried over  $\text{Na}_2\text{SO}_4$ , filtered, and evaporated under *vacuum* to afford product **39** a yellow oil (276 mg, 50 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 8.42 (s, 1H,  $\text{H}_4$  aryl), 8.15 (d, 1H,  $J=8.0$  Hz,  $\text{H}_2$  aryl), 8.04 (d, 1H,  $J=8.0$  Hz,  $\text{H}_6$  aryl), 7.62 (t, 1H,  $J=8.0$  Hz,  $\text{H}_1$  aryl), 5.94 (m, 1H,  $\text{H}_{13}$  NH); 5.09 (m, 1H,  $\text{H}_{16}$  NH); 3.23 (dd, 2H,  $J_1=8.0$  Hz,  $J_2=4.0$  Hz,  $\text{H}_{14}$   $\text{CH}_2$ ); 3.08 (dd, 2H,  $J_1=8.0$  Hz,  $J_2=4.0$  Hz,  $\text{H}_{15}$   $\text{CH}_2$ ); 2.65 (q, 2H,  $J=8.0$  Hz,  $\text{H}_9$   $\text{CH}_2$ ); 1.39 (s, 9H,  $\text{H}_{21}\text{-H}_{22}\text{-H}_{23}$   $\text{CH}_3$ ); 1.30 (t, 3H,  $J=8.0$  Hz,  $\text{H}_{10}$   $\text{CH}_3$ ).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.84; 28.30 (3C); 31.87; 40.01; 44.35; 55.04; 126.43; 129.65; 130.88; 131.75; 137.67; 140.86; 172.44; 199.34.

ESI-MS positive mode calculated for  $C_{16}H_{24}O_5N_2S_1$  356.1406, found 379.1306 ([M+Na<sup>+</sup>]), 735.2714 ([2M+Na<sup>+</sup>]).

**Ethyl 4-(3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-3-methyl-2,4-dioxobutanoate (40)**

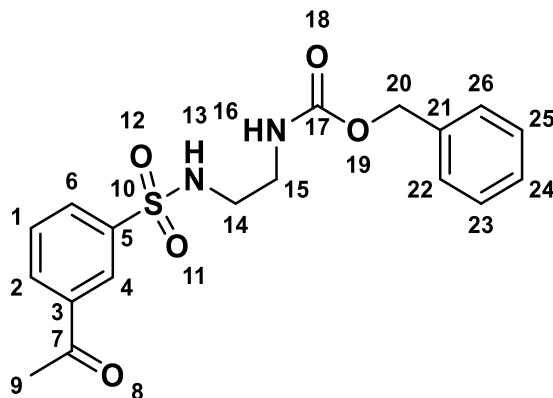


The same protocol as for the synthesis of the product **31** was applied to the ketone **39**.<sup>233</sup> The title compound **40** was collected as a yellow oil (205 mg, estimated 25 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 8.49 (s, 1H, H<sub>4</sub> aryl); 8.20 (d, 1H,  $J$ =8.0 Hz, H<sub>2</sub> aryl); 8.13 (d, 1H,  $J$ =8.0 Hz, H<sub>6</sub> aryl); 7.71 (t, 1H,  $J$ =8.0 Hz, H<sub>1</sub> aryl); 5.57 (m, 1H, H<sub>20</sub> NH); 5.08 (q,  $J$ =8.0 Hz, 1H, H<sub>10</sub>); 4.92 (m, 1H, H<sub>23</sub> NH); 4.31 (q, 2H,  $J$ =8.0 Hz, H<sub>16</sub> CH<sub>2</sub>); 3.26 (dd, 2H,  $J_1$ =4.0 Hz,  $J_2$ =8.0 Hz, H<sub>21</sub> CH<sub>2</sub>); 3.15 (dd, 2H,  $J_1$ =4.0 Hz,  $J_2$ =8.0 Hz, H<sub>22</sub> CH<sub>2</sub>); 1.49 (d, 3H,  $J$ =8.0 Hz, H<sub>31</sub> CH<sub>3</sub>); 1.45 (s, 9H, H<sub>28</sub>-H<sub>29</sub>-H<sub>30</sub> CH<sub>3</sub>); 1.34 (t, 3H,  $J$ =8.0 Hz, H<sub>17</sub> CH<sub>3</sub>).

ESI-MS positive mode calculated for  $C_{20}H_{28}O_8N_2S_1$  456.1566, found 479.1472 ([M+Na<sup>+</sup>]) and 935.3010 ([2M+Na<sup>+</sup>]).

**Benzyl 2-((3-acetylphenyl)sulfonamido)ethyl)carbamate (41)**

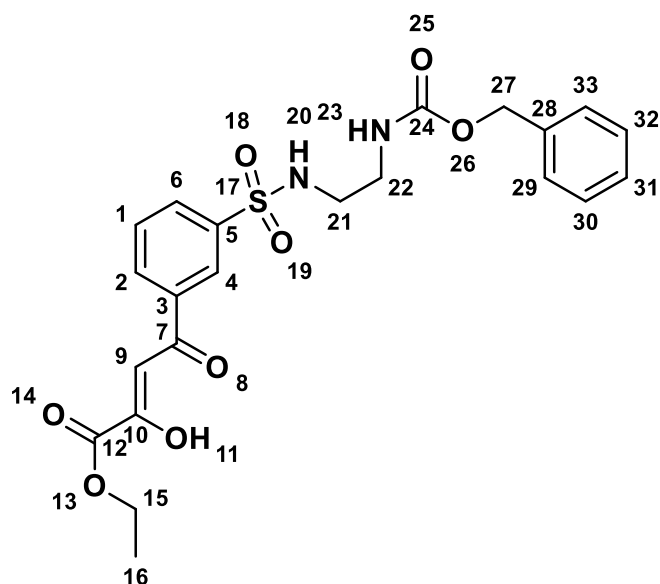


This compound was synthesised by adapting a previously described protocol.<sup>258</sup> To a solution of the deprotected ketone **37** (1 eq, 0.744 mmol, 253 mg) in a bi-phase diethyl ether (3 mL) and water (3.72 mL), was added Na<sub>2</sub>CO<sub>3</sub> (3 eq, 2.23 mmol, 237 mg). The solution was cooled on ice before the dropwise addition of benzylchloroformate (1 eq, 0.744 mmol, 0.106 mL). The solution was stirred at room temperature for 2.5h and monitored by TLC (ethyl acetate 80/20 hexane). The two phases were then separated, and the organic layer was wash with brine and water, dried over MgSO<sub>4</sub>, and filtered. The solvent was removed *in vacuum* and the resulting residue was purified by flash chromatography column on silica gel (ethyl acetate 80/20 hexane). The title compound **41** was collected as a yellow oil (195 mg, estimated 60 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 8.40 (s, 1H, H<sub>4</sub> aryl); 8.10 (d, *J*=8.0 Hz, 1H, H<sub>2</sub> aryl); 8.03 (d, *J*=8.0 Hz, 1H, H<sub>6</sub> aryl); 7.48 (t, *J*=8.0 Hz, 1H, H<sub>1</sub> aryl); 7.30 (m, 5H, H<sub>22-26</sub>); 6.04 (m, 1H, H<sub>13</sub> NH); 5.54 (m, 1H, H<sub>16</sub> NH); 5.03 (s, 2H, H<sub>20</sub>); 3.28 (m, 2H, H<sub>14</sub> CH<sub>2</sub>); 3.08 (m, 2H, H<sub>15</sub> CH<sub>2</sub>); 2.60 (s, 3H, H<sub>9</sub>). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 26.70; 40.74; 43.31; 66.61; 125.83; 126.70; 127.97 (2C); 128.16; 128.53; 129.71; 131.10; 132.12; 136.24; 137.79; 140.8; 156.99; 196.80.

ESI-MS positive mode calculated for C<sub>18</sub>H<sub>20</sub>O<sub>5</sub>N<sub>2</sub>S<sub>1</sub> 376.1093, found 399.0981 ([M+Na<sup>+</sup>]) and 775.2032 ([2M+Na<sup>+</sup>]).

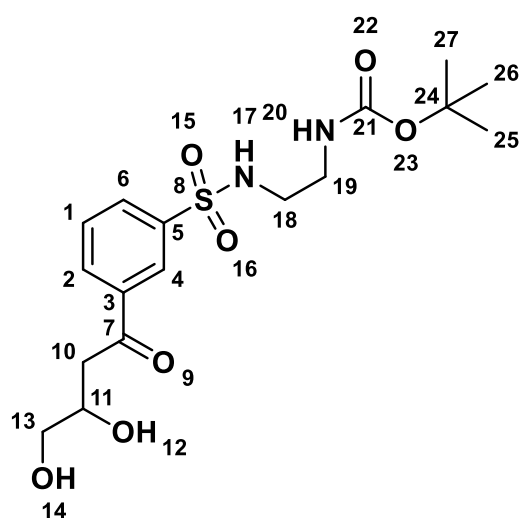
**Methyl4-(3-(N-(2-(((benzyloxy)carbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate (42)**



The same protocol as for the synthesis of the product **31** was applied to the Cbz-ketone **41**.<sup>233</sup> The title compound **42** was collected as a yellow oil (775 mg, estimated 69 %).

<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>Cl<sub>3</sub>): δ = 8.41 (s, 1H, H<sub>4</sub> aryl), 8.15 (d, 1H, H<sub>2</sub> aryl), 7.93 (d, 1H, H<sub>6</sub> aryl), 7.58 (t, 1H, H<sub>1</sub> aryl), 7.33 (m, 5H, H<sub>29-33</sub>); 6.80 (s, 1H, H<sub>9</sub>); 5.01 (s, 2H, H<sub>27</sub>); 4.32 (dd, 2H, J<sub>1</sub>=8.0 Hz, J<sub>2</sub>=16 Hz, H<sub>15</sub> CH<sub>2</sub>); 3.15 (m, 2H, H<sub>21</sub> CH<sub>2</sub>); 2.96 (m, 2H, H<sub>22</sub> CH<sub>2</sub>); 1.23 (s, 3H, H<sub>16</sub> CH<sub>3</sub>). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>Cl<sub>3</sub>): δ = 16.98; 40.37; 42.32; 45.90; 56.94; 60.17; 66.11; 125.50; 127.42 (2C); 128.06 (3C); 129.04 (3C); 130.72 (2C); 140.75; 157.44; 171.64; 218.50.

**tert-butyl(2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate (43)**



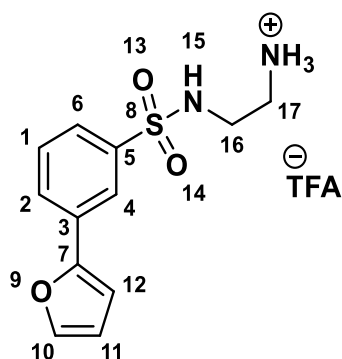
This compound was synthesised by adapting a previously described protocol.<sup>270</sup> To a cooled stirred mixture of the *N*-Boc-β-diketoester **31** (1 eq, 1.36 mmol, 601 mg) in anhydrous THF (44 mL) under nitrogen, was slowly added lithium aluminium hydride (4 eq, 5.44 mmol, 206 mg). The reaction mixture was stirred for 3h on ice and monitored by TLC (ethyl acetate 80/20 hexane). The reaction was treated with ethyl acetate and water at 0°C. The aqueous phase was extracted twice with ethyl acetate. The organic layers were washed with brine and water, dried over MgSO<sub>4</sub> and filtered. The solvent was removed *in vacuum* and the resulting residue was purified by flash chromatography column on silica gel (ethyl acetate 50/50 hexane). The title compound **43** was collected as a yellow oil (220 mg, estimated 36 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 8.47 (s, 1H, H<sub>4</sub> aryl); 8.19 (d, 1H, J=8.0 Hz, H<sub>2</sub> aryl); 8.09 (d, 1H, J=8.0 Hz, H<sub>6</sub> aryl); 7.66 (t, 1H, J=8.0 Hz, H<sub>1</sub> aryl); 5.75 (m, 1H, H<sub>17</sub> NH); 4.98 (m,

$^1\text{H}$  NMR ( $\text{H}_2\text{O}$  NH); 4.39 (m, 1H,  $\text{H}_7$  CH); 3.67 (m, 1H,  $\text{H}_{10}$   $\text{CH}_2$ ); 3.51 (m, 1H,  $\text{H}_{10}$   $\text{CH}_2$ ); 3.34 (m, 2H,  $\text{H}_{13}$   $\text{CH}_2$ ); 3.30 (m, 2H,  $\text{H}_{18}$   $\text{CH}_2$ ); 3.11 (m, 2H,  $\text{H}_{19}$   $\text{CH}_2$ ); 1.45 (s, 9H,  $\text{H}_{25}$ - $\text{H}_{26}$ - $\text{H}_{27}$   $\text{CH}_3$ ).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 28.33 (3C); 40.24; 41.89; 43.41; 65.86; 68.45; 79.88; 124.76; 126.77; 129.68; 131.41; 137.64; 140.84; 156.58; 207.21.

ESI-MS positive mode calculated for  $\text{C}_{17}\text{H}_{26}\text{O}_7\text{N}_2\text{S}_1$  402.1461, found 425.1353 ( $[\text{M}^+ + \text{Na}^+]$ ) and 827.2820 ( $[2\text{M}^+ + \text{Na}^+]$ ).

**N-(2-aminoethyl)-3-(furan-2-yl)benzenesulfonamide (44)**

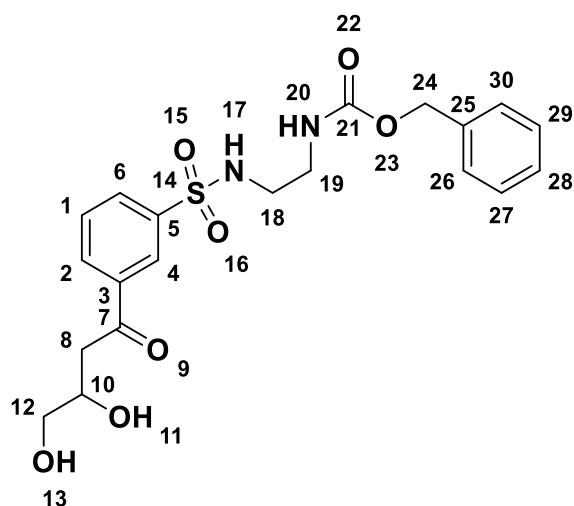


To a cooled solution of the ketodiol **43** (1 eq, 0.54 mmol, 220 mg) in dichloromethane (0.72 mL), was added dropwise under stirring trifluoroacetic acid (11 eq, 5.99 mmol, 0.46 mL). The solution was then stirred overnight at room temperature and monitored by TLC (ethyl acetate 60/40 hexane). The solution was evaporated under *vacuum* and co-distilled three times with dichloromethane to obtain product **44** as a yellow oil (134 mg, estimated 70 %).<sup>250</sup>

$^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  = 8.19 (s, 1H,  $\text{H}_4$  aryl); 7.98 (d, 1H,  $J=8.0$  Hz,  $\text{H}_2$  aryl); 7.77 (d, 1H,  $J=8.0$  Hz,  $\text{H}_6$  aryl); 7.65 (t, 1H,  $J=8.0$  Hz,  $\text{H}_1$  aryl); 7.64 (m, 1H,  $\text{H}_{11}$  CH); 6.96 (m, 1H,  $\text{H}_{10}$  CH); 6.59 (m, 1H,  $\text{H}_9$  CH); 3.13 (m, 2H,  $\text{H}_{16}$   $\text{CH}_2$ ); 3.09 (m, 2H,  $\text{H}_{17}$   $\text{CH}_2$ ).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  = 39.28; 40.00; 106.74; 111.76; 121.40; 125.07; 127.39; 129.64; 132.10; 140.29; 143.21; 151.94.

ESI-MS positive mode calculated for  $\text{C}_{12}\text{H}_{15}\text{O}_3\text{N}_2\text{S}_1^+$  266.0798, found 267.0813 ( $[\text{M} + \text{H}^+]$ ).

**Benzyl(2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate (45)**

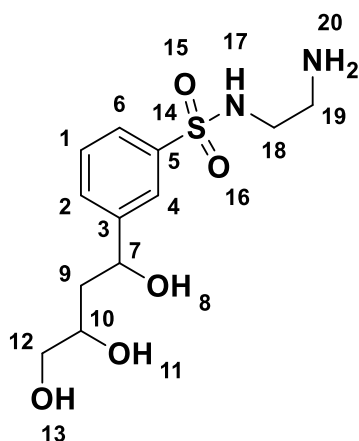


The same protocol as for the synthesis of the product **43** was applied to the Cbz-  $\beta$ -diketoester **42**.<sup>270</sup> The title compound **45** was collected as a yellow oil (172 mg, estimated 15 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 8.41 (s, 1H,  $\text{H}_4$  aryl); 8.06 (d, 1H,  $J=8.0$  Hz,  $\text{H}_2$  aryl); 7.99 (d, 1H,  $J=8.0$  Hz,  $\text{H}_6$  aryl); 7.51 (t, 1H,  $J=8.0$  Hz,  $\text{H}_1$  aryl); 7.28 (m, 5H,  $\text{H}_{26-30}$ ); 6.31 (m, 1H,  $\text{H}_{17}$  NH); 5.66 (m, 1H,  $\text{H}_{20}$  NH); 4.98 (s, 2H,  $\text{H}_{24}$ ,  $\text{CH}_2$ ); 4.30 (m, 1H,  $\text{H}_{10}$  CH); 3.66 (m, 1H,  $\text{H}_8$   $\text{CH}_2$ ); 3.57 (m, 1H,  $\text{H}_8$   $\text{CH}_2$ ); 3.22 (m, 2H,  $\text{H}_{18}$   $\text{CH}_2$ ); 3.10 (m, 2H,  $\text{H}_{12}$   $\text{CH}_2$ ); 3.04 (m, 2H,  $\text{H}_{19}$   $\text{CH}_2$ ).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 40.70; 41.83; 43.10; 65.84; 66.87; 68.43; 124.77; 126.69; 127.96 (2C); 128.18; 128.53; 248.53; 129.69; 131.37; 136.20; 137.58; 140.72; 157.01; 198.56.

ESI-MS positive mode calculated for  $\text{C}_{20}\text{H}_{24}\text{O}_7\text{N}_2\text{S}_1$  436.1304, found 459.1196 ( $[\text{M}^+ + \text{Na}^+]$ ) and 895.2459 ( $[2\text{M}^+ + \text{Na}^+]$ ).

**N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide (46)**



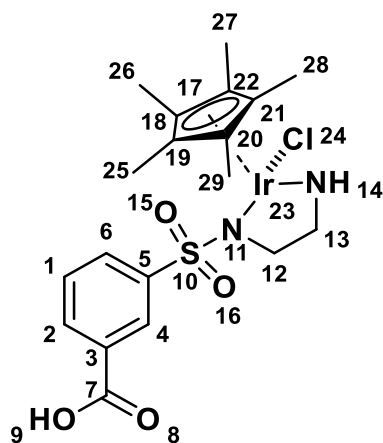
This compound was synthesised by adapting a previously described protocol.<sup>259</sup> To a Pd/C (1.3 eq, 46 mg, 0.43 mmol) powder purged under N<sub>2</sub>, was added the Cbz-ketodiol **45** (1 eq, 0.39 mmol, 172 mg) in methanol (6 mL). The solution was then flushed with H<sub>2</sub> gas and stirred under H<sub>2</sub> atmosphere (balloon) at room temperature for 24h.

The suspension was filtrated through a celite pad and washed with MeOH (2 x 5 mL). The filtrate was removed *in vacuo* to afford the title compound **46** as a yellow oil (90 mg, estimated 74 %).

<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ = 7.96 (s, 1H, H<sub>4</sub> aryl); 7.81 (t, 1H, *J*=8.0 Hz, H<sub>2</sub> aryl); 7.71 (t, 1H, *J*=8.0 Hz, H<sub>6</sub> aryl); 7.51 (m, 1H, H<sub>1</sub> aryl); 3.96 (m, 1H, H<sub>10</sub>); 3.62 (m, 1H, H<sub>7</sub> CH); 3.53 (d, 2H, *J*=4.0 Hz, H<sub>12</sub> CH<sub>2</sub>); 3.13 (m, 2H, H<sub>18</sub> CH<sub>2</sub>); 3.05 (m, 2H, H<sub>19</sub> CH<sub>2</sub>); 1.83 (m, 1H, H<sub>9</sub> CH<sub>2</sub>); 1.74 (m, 1H, H<sub>9</sub> CH<sub>2</sub>). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD): δ = 39.50; 40.52; 41.97; 65.79; 68.69; 69.97; 124.36; 125.64; 129.12; 130.50; 139.46; 139.51.

ESI-MS positive mode calculated for C<sub>12</sub>H<sub>20</sub>O<sub>5</sub>N<sub>2</sub>S<sub>1</sub> 304.1093, found 305.1180 ([M+H<sup>+</sup>]).

**[IrClCp\*]-(4-carboxyphenylsulfonyl)-ethylenediamine acid (47)**



This compound was synthesised by adapting a previously described protocol.<sup>58</sup> To a suspension of pentamethylcyclopentadienyl iridium(III) chloride dimer (1 eq, 0.063 mmol, 50 mg) in MeOH (2.3 mL) was added the deprotected acid **36** (3 eq, 0.19 mmol, 46 mg) in a MeOH (1 mL) and water (1 drop) solution. Triethylamine (4 eq) was then added dropwise until the solution turns a clear yellow. The solution was then stirred for 1h and the solvent was removed under vacuum. The resulting residue was purified

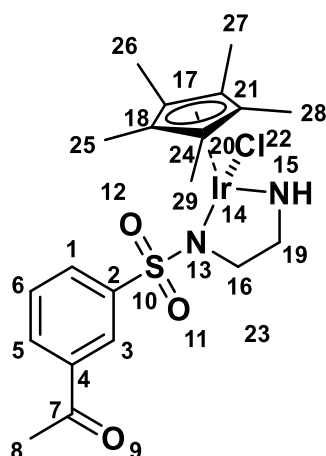


by flash chromatography column on silica gel (DCM 98/2 MeOH to 100% MeOH). The title compound **47** was collected as a yellow powder (21 mg, 55 %).

$^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  = 8.53 (s, 1H, H<sub>4</sub> aryl); 8.19 (d, 1H,  $J$ =8.0 Hz, H<sub>2</sub> aryl); 7.90 (d, 1H,  $J$ =8.0 Hz, H<sub>6</sub> aryl); 7.40 (t, 1H,  $J$ =8.0 Hz, H<sub>1</sub> aryl); 2.74 (m, 2H, H<sub>12</sub> CH<sub>2</sub>); 2.63 (m, 2H, H<sub>13</sub> CH<sub>2</sub>); 1.82 (s, 15H, Cp\*).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  = 7.99 (5C); 47.39; 48.03; 85.37 (5C); 127.74; 128.66; 130.02; 130.76; 131.05; 136.00; 178.45.

ESI-MS positive mode calculated for C<sub>19</sub>H<sub>25</sub>IrClO<sub>4</sub>N<sub>2</sub>S<sub>1</sub> 605.0853, found 605.1200 ([M]), 569.1490 ([M+H- $^{37}\text{Cl}$ ]) and 570.1483 ([M- $^{35}\text{Cl}$ ]).

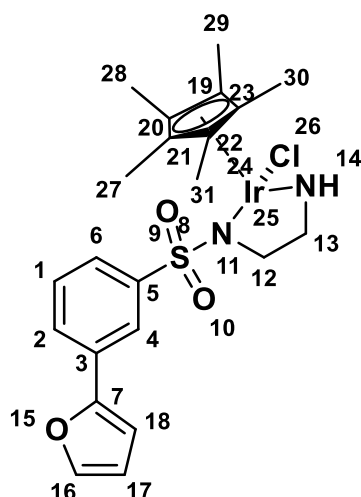
**[IrClCp\*]-3-acetyl-N-(2-aminoethyl)benzenesulfonamide (48)**



This compound was synthesised by adapting a previously described protocol.<sup>58</sup> To a suspension of pentamethylcyclopentadienyl iridium(III) chloride dimer (1 eq, 0.25 mmol, 148 mg) in MeOH (3 mL) under N<sub>2</sub> atmosphere was added a solution of the deprotected ketone **37** (2 eq, 121 mg, 0.5 mmol) in MeOH (3 mL). Triethylamine (4 eq) was then added dropwise until completion of the reaction (yellow solution). The solution was evaporated under *vacuum* and dissolved again in DCM (1 mL). Hexane was added for precipitation and the resulting suspension was filtered to afford product **48** as a yellow powder (30 mg, 20 %).

$^1\text{H}$  NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 8.55 (s, 1H, H<sub>3</sub> aryl); 8.11 (dt, 1H,  $J$ =8.0 Hz, H<sub>5</sub> aryl); 7.98 (dt, 1H,  $J$ =8.0 Hz, H<sub>1</sub> aryl); 7.45 (t, 1H,  $J$ =8.0 Hz, H<sub>6</sub> aryl); 2.72 (m, 2H, H<sub>16</sub> CH<sub>2</sub>); 2.67 (s, 3H, H<sub>8</sub> CH<sub>3</sub>); 2.64 (m, 2H, H<sub>19</sub> CH<sub>2</sub>); 1.77 (s, 15H, Cp\*).

**[IrClCp\*]-N-(2-aminoethyl)-3-(furan-2-yl)benzenesulfonamide (49)**



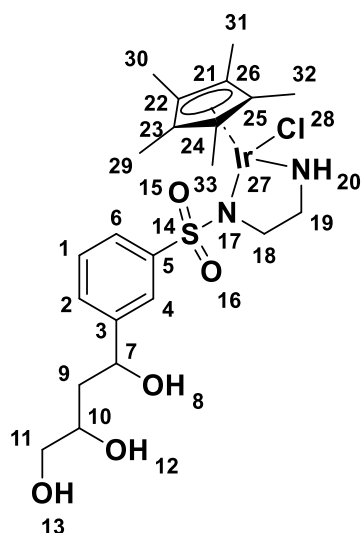
The same protocol as for the synthesis of the product **48** was applied to the furan **44**.

<sup>58</sup> The title compound **49** was collected as a red powder (29 mg, 43 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 8.29 (s, 1H, H<sub>4</sub> aryl); 7.82 (d, 1H, *J*=8.0 Hz, H<sub>2</sub> aryl); 7.69 (d, 1H, *J*=8.0 Hz, H<sub>6</sub> aryl); 7.51 (m, 1H, H<sub>16</sub> CH); 7.46 (t, 1H, *J*=8.0 Hz, H<sub>1</sub> aryl); 6.80 (m, 1H, H<sub>17</sub> CH); 6.47 (m, 1H, H<sub>18</sub> CH); 2.74 (m, 2H, H<sub>12</sub> CH<sub>2</sub>); 2.64 (m, 2H, H<sub>13</sub> CH<sub>2</sub>); 1.78 (s, 15H, Cp\*). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 9.42 (5C); 48.52; 50.61; 85.49 (5C); 105.95; 111.79; 123.45; 125.33; 126.74; 128.48; 130.73; 142.13; 143.75; 153.47.

ESI-MS positive mode calculated for  $C_{22}H_{27}ClIrO_3N_2S_1$  627.1060, found 591.1412 ( $[M^{+37}Cl]$ ), 593.1451 ( $[M^{+35}Cl]$ ), 629.1197 ( $[M+2H^+]$ ) and 651.1018 ( $[M+H^+Na]$ ).

**[IrClCp\*]-N-(2-aminoethyl)-3-(3,4dihydroxybutanoyl)benzenesulfonamide (50)**



The same protocol as for the synthesis of the product **48** was applied to the triol **46**.<sup>58</sup>

The title compound **50** was collected as an orange powder (43 mg, 43 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  = 8.02 (s, 1H,  $\text{H}_4$  aryl); 7.85 (d, 1H,  $J=8.0$  Hz,  $\text{H}_2$  aryl); 7.49 (d, 1H,  $J=8.0$  Hz,  $\text{H}_6$  aryl); 7.40 (t, 1H,  $J=8.0$  Hz,  $\text{H}_1$  aryl); 3.92 (m, 1H,  $\text{H}_{10}$ ); 3.60 (m, 1H,  $\text{H}_7$  CH); 3.52 (m, 2H,  $\text{H}_{11}$   $\text{CH}_2$ ); 2.60 (m, 2H,  $\text{H}_{18}$   $\text{CH}_2$ ); 2.50 (m, 2H,  $\text{H}_{19}$   $\text{CH}_2$ ); 1.92 (m, 2H,  $\text{H}_9$   $\text{CH}_2$ ); 1.76 (s, 15H,  $\text{Cp}^*$ ).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  = 8.04 (5C); 41.80; 42.44; 48.03; 65.99; 66.34; 71.70; 85.41 (5C); 125.14; 125.60; 126.33; 126.59; 143.07; 145.06.

ESI-MS positive mode calculated for  $\text{C}_{22}\text{H}_{33}\text{ClIrO}_5\text{N}_2\text{S}_1$  665.1428, found 629.1779 ( $[\text{M}^+ - ^{37}\text{Cl}]$ ), 631.1793 ( $[\text{M}^+ - ^{35}\text{Cl}]$ ) and 667.1558 ( $[\text{M} + \text{H}^+]$ ).

## 5 Enzyme engineering: non-covalent conjugation and catalysis of imine reduction

### 5.1 Introduction

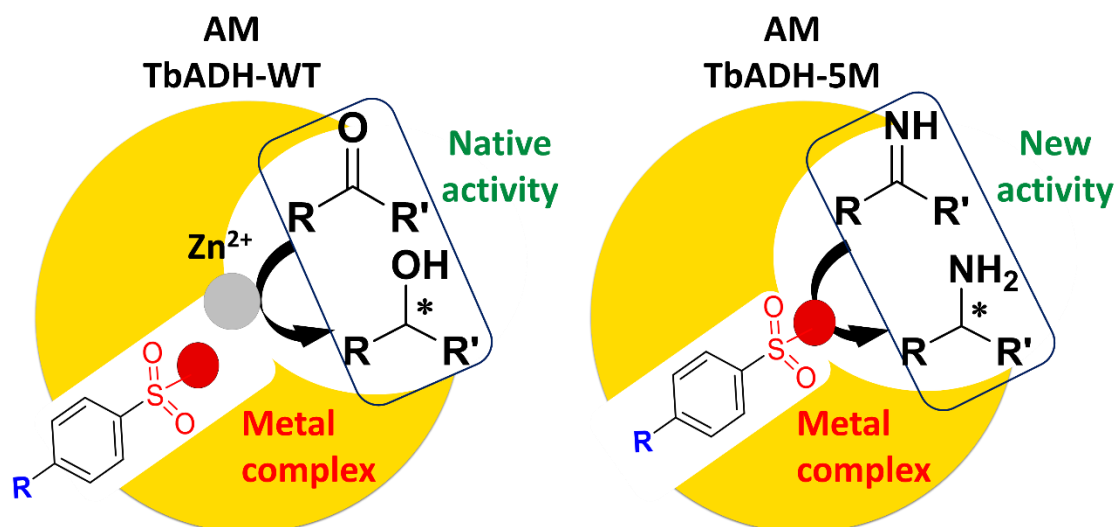
The focus of this thesis is to design, synthesise, and test nicotinamide cofactor structural analogues that bind strongly inside three selected ADHs. Following the synthesis of three transition metal catalysts based on *in silico* study, this chapter will investigate the supramolecular binding of the anchor structures and the entire metal complexes inside three alcohol dehydrogenases (ADHs). Finally, the biocatalytic reduction of cyclic imine and the stereoselectivity of the resulting artificial metalloenzymes (AMs) will be investigated.

The bioconjugation was investigated by means of isothermal titration calorimetry (ITC) and inhibition tests against the catalysis with natural ADHs cofactors. This gave insight into the affinity between the ADHs and the catalysts. The catalytic activity of both AMs and the metal catalysts alone was subsequently analysed with HPLC analysis in aqueous media.

In this study, ADHs have been selected because of their similarities with imine reductases, which naturally catalyse the targeted reaction. ADHs use the nicotinamide cofactor NAD(P)H to catalyse a hydride transfer reaction with the carbon double bonded to oxygen. With a similar C=X reduction reaction and the dependence on the same cofactors, it was hypothesised that the well-known ADHs were of suitable scaffold shape to engineer AMs toward the reduction of double bonded carbon-heteroatom. Besides, ADHs presents better characteristics than imine reductases to fit in a synthesis. For example, ADHs show higher tolerance to solvents and heat, and a wider substrate scope is accepted.<sup>287</sup>

Three ADHs genes, available in our laboratory, were chosen to build the AMs: Horse Liver ADH possessing an N-terminal hexahistidine tag (His-HLADH),<sup>288</sup> *Thermoanaerobacter Brockii* ADH wild type (TbADH-WT)<sup>289</sup> and a mutant of TbADH (TbADH-5M).<sup>109</sup> The latter enzyme variant contains five mutations of residues to alanine or serine (H59A-D150A-C203S-C283A-C295A), and is devoid of the catalytic

zinc ion, by the loss of two coordinating residues (H59 and D150). As already discussed in chapter 3, the space freed by  $\text{Zn}^{2+}$  favours the accommodation of the synthetic iridium catalysts, resulting in a binding closer to the substrate (**Figure 72**).



**Figure 72:** Schematic representation of AMs with TbADH-WT scaffold (left) and after site-directed mutagenesis with TbADH-5M lacking the catalytic  $\text{Zn}^{2+}$  (right). The loss of the metal ion freed space for the transition metal catalyst.

The two wild type enzymes and the zinc-devoid variant will be used as models for the non-covalent binding of transition metal catalyst to engineer new AMs.

## 5.2 Materials and Methods

### 5.2.1 Reagents and materials

All reagents were purchased from Sigma-Aldrich, *E. coli* BL21 (DE3) competent cells were obtained from Merck and the enzymes were purchased from New England Biolabs, unless otherwise stated in the text (Appendix 2). Sterilisation of all media and flasks was performed with a Prestige Classic 2100 benchtop autoclave (Medstore Medical). The cultures were grown in 2 L Erlenmeyer flasks, in a New Brunswick Scientific Innova 40 Incubator Shaker. All cells were harvested using an Avanti J-26 XP centrifuge (Beckman Coulter) and a 5810R centrifuge (Eppendorf). One Shot Cell Disruptor (Constant System Ltd.) was used to lyse the cells. An AKTA Fast Liquid Protein Chromatography System (GE Healthcare) was used to purify the proteins.

Kinetic assays were performed with an Apollo Scientific Ltd (UK). UV mini 1240 UV-Vis spectrophotometer (Schimadzu), equipped with CPS-100 temperature controller. ITC was assessed using a Microcal iTC200 titration calorimeter (MicroCal Inc.,

Northampton, MA). For catalytic tests, an HPLC Agilent 1220 Infinity was used with a Chiracel OD 4.6 mm x 250 mm. PS 10  $\mu$ m column.

## 5.2.2 *Methods*

### 5.2.2.1 Enzymes overexpression and purification

#### - ***Growth media***

Media components were sterilised by autoclaving (121 °C, 20 min, 15 psi) and mixed when cooled. Media solutions were prepared with Lysogeny broth (LB) media in deionised water and stored at room temperature. Where solid medium was required, LB agar (1.5 % w/v, Fisher scientific) was added to the medium prior to autoclaving. Where applicable, antibiotics were added to growth media. Antibiotics were prepared as 1000 times concentrate of the working concentration. Carbenicillin (50 mg/mL) solutions were filter sterilised and stored at -20 °C. All cultivations were performed at 37 °C, 200 rpm and in baffled Erlenmeyer flasks, with nominal volumes five times the culture volume unless otherwise stated.

#### - ***His-HLADH***

Commercially available chemically competent *E. coli* BL21 (DE3) cells (Merck) were transformed with the plasmid pRSETb-HisHLADH, according to the manufacturer's instructions. Briefly, the plasmid (3  $\mu$ L) was added to the cells (100  $\mu$ L) and incubated on ice for 30 min. A heat shock of 42 °C for 10 seconds was applied and the cells were cooled on ice for 5 min. SOC medium (950  $\mu$ L) was then added, and the cells were shaking for 60 min at 37 °C. The resulting culture was plated and incubated overnight at 37 °C.

Transformed cells were plated on LB agar supplemented with carbenicillin and incubated for 16 h at 37 °C. A single colony was subsequently used to inoculate 100 mL of LB medium supplemented with carbenicillin, and incubated for 7h at 37 °C, 180 rpm before cryostocks of cells were prepared and stored at -80 °C in 20 % glycerol. An appropriate volume of this starter culture was then used to inoculate fresh LB medium (800 mL) supplemented with carbenicillin, in an Erlenmeyer flask (total volume 2 L) to an OD<sub>600</sub> of 0.1-0.2. The resulting culture was then incubated at 37 °C, 200 rpm and until the OD<sub>600</sub> reached 0.6-0.8. Protein expression was induced by the

addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG; 1 M stock sterilised by filtration; 0.4 mM final concentration). Induced cultures were incubated for 16 h at 30 °C and 200 rpm, after which the cells were harvested by centrifugation (5 000 g, 15 min, 4 °C) and the pellets were stored at -80 °C prior to purification.

Cells pellets were then re-suspended in 15 mL buffer A (**Table 13**) at pH 7.39, supplemented with 0.35 mg/mL lysozyme from chicken egg. The suspension was stirred on ice until homogenous (30 min), then sonicated for 5 minutes (35 % amplitude) in 10 s bursts, with 30 s intervals for cooling. Cell debris were removed by centrifugation at 38 000 g at 4°C for 30 min. The cell free extract was filtered through 0.22  $\mu$ m filter, then purified by affinity chromatography on a 5 mL His-Trap FF crude column pre-packed with Ni-Sepharose™ Fast Flow resin (GE Healthcare) and pre-charged with NiSO<sub>4</sub> (0.1 M), with a flow rate of 2 mL.min<sup>-1</sup>. The column was pre-equilibrated with 10 CV (column volumes) of the buffer A. The protein was eluted in 2.5 mL fractions by the addition of buffer B (**Table 13**). Elution fractions displaying a strong A280 nm absorbance were analysed by SDS-PAGE and pooled together. EDTA was removed from the pooled fractions by dialysis by exchange into buffer C (**Table 13**). The protein was stored at -20 °C.

**Table 13:** Buffers used for His-HLADH purification. Buffer A corresponded to the binding buffer, buffer B was the elution buffer and buffer C the stock buffer.

Buffer A	20 mM Tris, 500 mM NaCl, 20 mM imidazole
Buffer B	20 mM TrisHCl, 500 mM NaCl, 15 mM EDTA
Buffer C	100 mM Tris

#### - *TbADH-WT and TbADH-5M*

Commercially available chemically competent *E. coli* BL21 (DE3) cells were transformed with the plasmid pET-21a(+):TbADH, according to the manufacturer's instructions. The previously described protocol was applied (induction with 1 mM IPTG and 1 mM ZnCl<sub>2</sub>). The pellets were also stored at -80 °C prior to purification.

Cells were suspended in buffer A (**Table 14**), supplemented with 1 mg/mL lysozyme, complete™ protease inhibitor cocktail (1/4<sup>th</sup> of tablet/10 mL) and Benzonase® endonuclease (1  $\mu$ L / mL). The suspension was stirred for 1 hour on ice, then sonicated for 5 minutes (35 % amplitude) in 10 s bursts, with 30 s intervals for cooling. Cell debris were removed by centrifugation (10 °C, 15 min, 25 000 g). The cell free

extract was purified by heat treatment at 60 °C for 15 minutes followed by incubation on ice for 10 minutes. This was followed by centrifugation (4 °C, 15 min, 16 900 g) to remove aggregated protein.

The cell free extract was filtered through a 0.22 µm filter, then purified by affinity chromatography on a 5 mL StrepTrap™ HP column (GE Healthcare) with a flow rate of 3 mL.min<sup>-1</sup>. The column was pre-equilibrated with 10 CV of the buffer A. The protein was eluted in 2.5 mL fractions by the addition of buffer B (**Table 14**). Elution fractions displaying a strong A280 nm absorbance were analysed by SDS-PAGE and pooled together. D-Desthiobiotin was removed from the pooled fractions by using a 10000 MWCO Viva-Spin 6 column (Sartorius), by exchange into buffer C (**Table 14**). The protein was stored at -20 °C.

**Table 14:** Buffers used for TbADH-WT and TbADH-5M purification. Buffer A corresponded to the binding buffer, buffer B was the elution buffer and buffer C the stock buffer.

Buffer A	100 mM Tris, 150 mM NaCl
Buffer B	100 mM TrisHCl, 150 mM NaCl, 2.5 mM desthiobiotin
Buffer C	100 mM Tris

#### - **Protein analysis**

Purity was assessed by SDS-PAGE. Protein samples (20 µL) were mixed with 20 µL of Laemmli sample buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromphenol blue and 0.125 M Tris HCl, pH approx. 6.8) and incubated at 95 °C for 5 minutes. PageRuler™ Plus Unstained Protein Ladder 10 to 250 kDa (5 µL) and samples (10 µL) were loaded onto an 8% polyacrylamide pre-cast gel (Bio-Rad) and ran at 200 V, 400 mA for 35 min, with Tris-Glycine-SDS running buffer (14.4 g/L glycine, 3 g/L Tris base, 1 g/L SDS). Gels were stained with InstantBlue (Expedeon).

Protein concentration was determined by the Bradford assay using a set of bovine serum albumin (BSA) standards, prepared via serial dilution in a range of 0.125 to 1 mg/mL, in addition to several dilutions of TbADH sample. Assays were performed in 96-well plates in triplicate, by mixing 300 µL sample Bradford reagent (Bio-Rad Laboratories) with 10 µL of either TbADH sample, BSA standard, or sample buffer (control). Samples and standards were incubated in the dark at room temperature for 30 minutes prior to measurement of the absorbance at 595 nm (FLUOstar

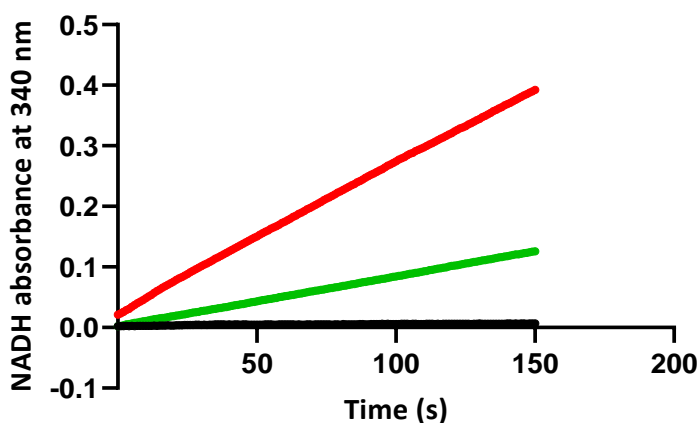


OPTIMA, BMG Labtech). The absorbances of BSA standards were used to generate a standard curve from which TbADH sample concentrations were estimated.

#### 5.2.2.2 Activity assays<sup>288,290</sup>

The NAD(P)<sup>+</sup> oxidation and NAD(P)H reduction by purified His-HLADH and TbADH-WT respectively, were monitored by following a previously reported method, via change in absorbance at 340 nm using a UV-spectrophotometer in kinetics mode, in quartz cuvettes with a pathlength of 1 cm. For all activity assays, controls were performed where the enzyme or the cofactor were replaced by the assay buffer, and minimal to no increase in absorbance was observed.

The rates of the enzymatic reactions were calculated using the slope of the curve of absorbance vs time in the first 60 s and corrected for background activity (in the absence of enzyme).



**Figure 73:** Example of initial rate of reaction for an enzymatic assay in the oxidative direction. NAD(P)H absorbance was monitored at 340 nm, with butan-2-ol and ethanol as the substrates: -●- TbADH-WT 150 mM butan-2-ol -●- His-HLADH 200 mM ethanol -●- Control.

The slope for the initial rate of reaction was converted to NAD(P)H concentration using the Beer-Lambert law:

$$A = \epsilon \cdot l \cdot c \quad (3)$$

where A = absorbance observed at 340 nm;  $\epsilon$  = extinction coefficient (NAD(P)H  $\epsilon_{340} = 6220 \text{ M}^{-1} \text{ cm}^{-1} = 0.00622 \text{ } \mu\text{M}^{-1} \text{ cm}^{-1}$ ); l = path length (cm) and c = concentration of NAD(P)H ( $\mu\text{M}$ ).

Activity was defined as  $\mu\text{mol NAD(P)H min}^{-1}$ . Specific activity was defined as  $\mu\text{mol NAD(P)H min}^{-1} \text{ mg(enzyme)}^{-1}$ .

Experiments were performed in triplicate and the standard errors of mean activity values calculated. Conditions were as follows in 1 mL quartz cuvette:

**Table 15:** Final concentrations for the oxidative and reductive reactions of His-HLADH and TbADH-WT.

	Oxidation reaction		Reduction reaction	
	His-HLADH	TbADH-WT	His-HLADH	TbADH-WT
<b>Enzymes</b>	15 µg	1 µg	15 µg	1 µg
<b>Cofactors</b>	NAD <sup>+</sup> 0.5 mM	NADP <sup>+</sup> 0.5 mM	NADH 0.1 mM	NADPH 0.1 mM
<b>Substrates</b>	Ethanol 4 mM	Butan-2-ol 150 mM	Acetaldehyde 10 mM	Butan-2-one 10 mM
<b>Buffer</b>	100 mM TrisHCl pH 8 up to 1 mL	100 mM TrisHCl pH 8 up to 1 mL	100 mM TrisHCl pH 8 up to 1 mL	100 mM TrisHCl pH 8 up to 1 mL
<b>Condition</b>	25 °C	40 °C	25 °C	40 °C

#### 5.2.2.3 Michaelis-Menten kinetic behaviour of His-HLADH

The tests were performed in the oxidative direction. The enzymes, NAD<sup>+</sup> and ethanol stocks were prepared in 100 mM TrisHCl buffer pH 8 and added to a 1 mL cuvette:

**Table 16:** Final concentrations for oxidative Michaelis-Menten assay on His-HLADH.

	His-HLADH
<b>Enzyme</b>	15 µg
<b>NAD<sup>+</sup></b>	0.013 mM to 7mM
<b>Ethanol</b>	4 mM
<b>Buffer</b>	100 mM TrisHCl pH 8 up to 1 mL

The solution was incubated for 5 minutes at 25 °C and the reaction was initiated by the addition of the enzyme, followed by a quick reverse of the cuvette for mixing.

The increase of absorbance at 340 nm was monitored and the previously described activity calculations (section 5.2.2.2, page 164) were applied, to obtain the specific activity for each cofactor concentrations. GraphPad Prism 8 was used to determine the Michaelis-Menten parameters  $K_M$  and  $V_{max}$  by non-linear regression, with a narrow confidence interval fit of the parameters (the Appendix 3). All experiments were run in triplicates and corrected for background activity (in the absence of enzyme).

#### 5.2.2.4 Inhibition assays of His-HLADH and TbADH-WT by the three synthesised transition metal complexes

The procedures described for Michaelis-Menten kinetics were applied for inhibition kinetics of the oxidation reaction. The cofactor concentration was kept fixed, and inhibitors were added with concentrations described in **Table 17**. The synthesised iridium complexes and their unprotected diamine form were used as inhibitors: Benzoic acid **36**; [IrClCp\*]-benzoic acid **47**; Furan **44**; [IrClCp\*]-furan **49**; Triol **46**; [IrClCp\*]-triol **50**.

**Table 17:** Final concentrations for oxidative inhibition assays

	His-HLADH	TbADH-WT
<b>Enzymes</b>	15 µg	1.0 µg
<b>Cofactors</b>	NAD <sup>+</sup> 0.5 mM	NADP <sup>+</sup> 0.5 mM
<b>Substrate</b>	Ethanol 4 mM	Butan-2-ol 150 mM
<b>Buffer</b>	100 mM TrisHCl pH 8 up to 1 mL, at 25 °C	100 mM TrisHCl pH 8 up to 1 mL, at 40 °C
<b>Inhibitors</b>	0.001 mM to 150 mM	0.001 mM to 40 mM

#### 5.2.2.5 Catalysis assays

The following reaction protocol was adapted from a procedure established by Ward and co-workers.<sup>291</sup> The imine substrate used was 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline and the corresponding amine product was 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline. The previously synthesised iridium complexes were used as catalysts: [IrClCp\*]-benzoic acid **47**; [IrClCp\*]-furan **49**; [IrClCp\*]-triol **50**.

The commercially available imine and amine were prepared in 100 mM Phosphate-buffered saline (PBS) pH 7 at 10 mM. The iridium complexes and the enzymes were prepared in PBS buffer pH 7 at 2 mM concentration and diluted to 250 µM concentration. Sodium formate was prepared at a concentration of 1 M in 100 mM PBS pH 7 buffer.

The following reagents were added to a vial: Sodium formate (150 µL, final conc. 300 mM), imine substrate (200 µL, final conc. 4 mM), iridium complex and enzymes (100

$\mu\text{L}$ , final conc.  $50\ \mu\text{M}$ ) and  $100\ \text{mM}$  PBS buffer pH 7 up to  $500\ \mu\text{L}$ . The reaction mixture was vortexed and incubated at  $25\ ^\circ\text{C}$  for His-HLADH and  $40\ ^\circ\text{C}$  for TbADH-WT and TbADH-5M. After 24 h, the mixture was cooled down and quenched by  $10\ \text{M}$  NaOH ( $55\ \mu\text{L}$ ). The product was extracted with  $1\ \text{mL}$  ethyl acetate. The organic phase was separated and dried over  $\text{MgSO}_4$ . The dried organic phase was transferred in a clean vial and analysed by HPLC using a Chiracel OD column ( $1\ \mu\text{L}$  / injection, mobile phase hexane / isopropanol / diethylamine 95/5/0.1,  $25\ ^\circ\text{C}$ , flow rate of  $0.85\ \text{mL/min}$ ). Experiments were run in triplicates and controls were performed with no iridium complexes and no enzymes.

A calibration curve was developed to assess the formation of amine product and the consumption of imine substrate. The commercially available substrate and product were added to a sodium formate and PBS buffer at several concentrations:  $0.1\ \text{mM}$ ,  $1\ \text{mM}$ ,  $2\ \text{mM}$ ,  $3\ \text{mM}$ ,  $4\ \text{mM}$ . The mixture was adjusted with NaOH, extracted with ethyl acetate, and dried over  $\text{MgSO}_4$ . The dried organic phase was analysed by HPLC following the previous protocol (Appendix 5, **Figure 102**).

The kinetic results for the three iridium catalysts were expressed in turnover frequency (TOF,  $\text{min}^{-1}$ ), defined as  $\mu\text{mol}$  of imine consumed per  $\mu\text{mol}$  of catalyst per hour.

#### 5.2.2.6 Isothermal Titration Calorimetry

The sample cell of the calorimeter ( $300\ \mu\text{L}$ ) was filled with a phosphate buffer solution of the four enzymes: TbADH-WT ( $13\ \mu\text{M}$  to  $129\ \mu\text{M}$ ), TbADH-5M ( $99$ - $123\ \mu\text{M}$ ) and His-HLADH ( $100\ \mu\text{M}$ ).

The spinning syringe ( $750\ \text{rpm}$ ) was filled with the  $\text{NAD(P)}^+ / \text{NAD(P)H}$  cofactors, the benzoic acid **37** and the  $\beta$ -diketoester **31** in phosphate buffer solution ( $2100$ - $2400\ \mu\text{M}$ ). The syringe injected the solution into the cell at  $25\ ^\circ\text{C}$ . Injections of  $2.4\ \mu\text{L}$  were made 17 times for 4.8 seconds and every 180 seconds over a period of 51 min. The data points were then collected every 2 seconds.

Data analysis was carried out using the NITPIC<sup>292</sup> and SEDPHAT<sup>293</sup> software, where the area of the peaks is integrated versus the molar ratio of ligands to protein (example of raw data in **Figure 76**).

## 5.3 Results and discussion

### 5.3.1 *Production of functional enzymes: His-HLADH, TbADH-WT and TbADH-5M*

#### 5.3.1.1 Expression and purification

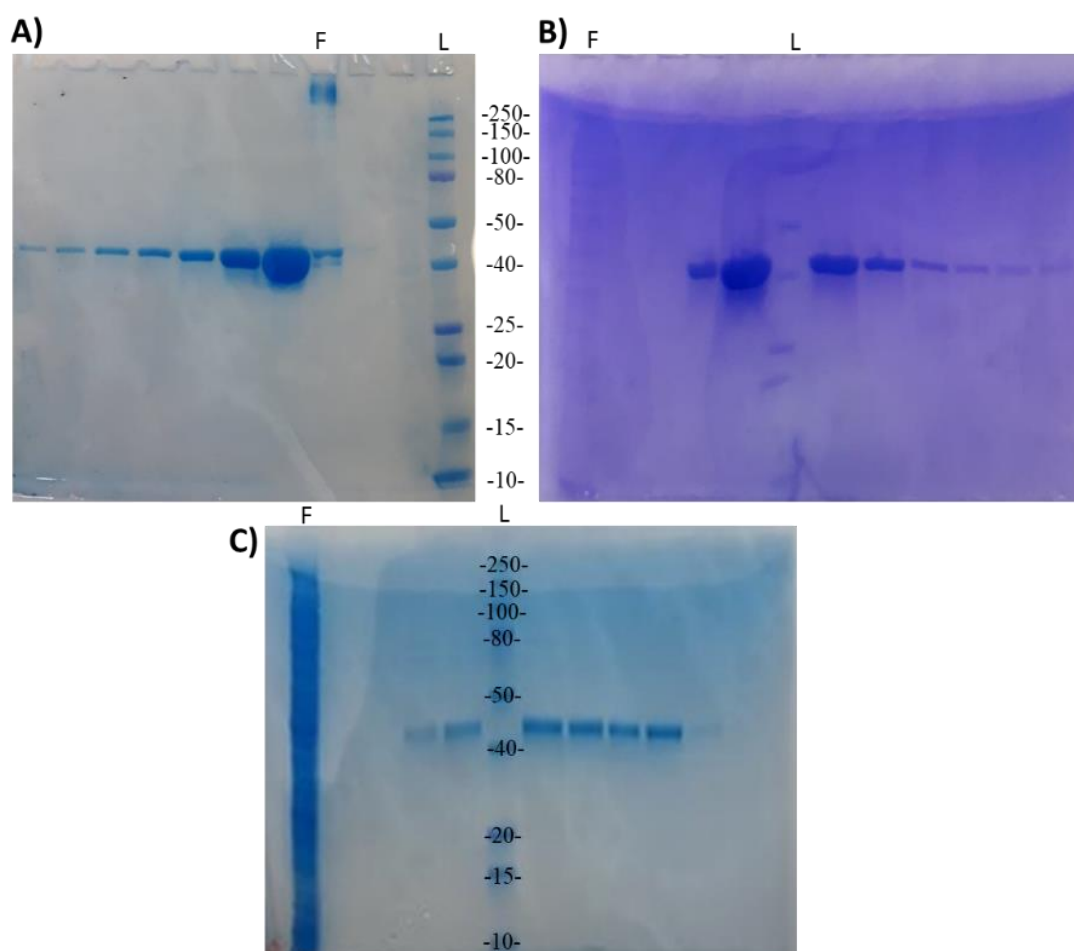
The plasmid for His-HLADH was obtained from Prof. Francesca Paradisi's group and the plasmid for TbADH-WT was synthesised by Biomatik (Ontario, Canada). Both genes were already available in the laboratory, into the expression vectors pRSETb for His-HLADH and pET-21a for TbADH-WT. The mutant TbADH-5M gene (5M-C37), already reported in previous work from our laboratory,<sup>289</sup> was available in the pET-21a(+) vector. The genes sequences are presented in Appendix 2.

Following previously described protocols (section 5.2.2.1, page 161), *E.coli* BL21 (DE3) was transformed with the three plasmids, and the three enzymes were expressed in LB medium with IPTG induction at 30 °C, prior to affinity chromatography purification.

Both TbADH variants contained a *N*-terminus Strep tag II (W-S-H-P-Q-F-E-K) and the HLADH contained a *C*-terminus hexahistidine tag (His-HLADH). The TbADHs proteins were purified successfully. However, for a successful and reproducible purification of His-HLADH, the existing protocol was optimised. Based on previous work in the laboratory, the first attempt using imidazole 500 mM to remove His-HLADH from the Ni<sup>2+</sup>- agarose matrix of the column resulted in the His-HLADH showing no activity for the oxidation of ethanol. These data may suggest an irreversible inhibition of His-HLADH by imidazole, as described by Dahl and McKinley-McKee.<sup>294</sup>

Based on the purification method developed by Paradisi and co-workers, EDTA was utilised as an alternative eluent buffer.<sup>288</sup> Following this protocol, the pooled His-HLADH fractions were dialysed overnight, as a precaution to remove EDTA known for its inhibitory effect on zinc-dependent enzymes. This second protocol proved to be successful, resulting in an enzyme activity consistent with literature (**Table 18**).

The purity of the enzymes was assessed by SDS-PAGE gel (**Figure 74**). The purified enzymes migrated at a single band of an apparent weight of ~44 kDa and ~40 kDa for TbADHs and His-HLADH respectively, corresponding to the expected ADHs subunit weight (Appendix 2).



**Figure 74:** Protein SDS-PAGE of three ADHs variants after purification of **A)** TbADH-WT, **B)** TbADH-5M and **C)** His-HLADH. F= flowthrough and L= Ladder.

The concentrations of the purified proteins were tested by the Bradford method and a total of 10 mg for His-HLADH, 17 mg TbADH-WT and 15 mg TbADH-5M for 400 mL of culture medium was found. The enzymes were then stored at -80 °C until further use.

#### 5.3.1.2 Determination of the specific activity

The activity of both His-HLADH and TbADH-WT was measured to confirm the integrity and functionality of the wild type enzymes. The specific activity was determined in the oxidative and the reductive direction by monitoring NAD(P)H spectroscopically at 340 nm. Ethanol and butan-2-ol were used with His-HLADH and TbADH-WT as substrate in the oxidation reactions, and acetaldehyde and butan-2-one were used for the reduction reactions. The activities were calculated following the published protocol described in section 5.2.2.2, page 164:

**Table 18:** Specific activity of His-HLADH and TbADH-WT for the oxidation and reduction of their natural substrates, compared to the literature reports.<sup>288,295</sup> One unit of activity is defined by the amount of enzyme that catalyses the formation of 1  $\mu$ mol of NAD(P)H per minute (U / mg). The protocol is described in section 5.2.2.2, page 164.

	Specific activity in this work	Literature references <sup>288,295</sup>
<b>His-HLADH<sup>a</sup></b>	Oxidation: 0.73 U / mg Reduction: 6.7 U / mg	Oxidation: 0.77 U / mg Reduction: no data
<b>TbADH-WT<sup>b</sup></b>	Oxidation: 62 U / mg Reduction: 9.3 U/mg	Oxidation: 78 U / mg Reduction: 7.6 U/mg

The experimental activities of both enzymes were considered in range with previously published data, confirming the enzymes integrity.

After the difficulties encountered for His-HLADH purification, steady state kinetic parameters were also determined for confirmation of a functional protein. The kinetic results in the oxidation of ethanol were  $K_M = 0.12$  mM (95 % CI  $0.07819 < K_M < 0.1773$ ) and  $V_{max} = 0.73$  U / mg, (95 % CI  $0.6625 < V_{max} < 0.7924$ ) compared to the reported values of  $K_M = 0.26$  mM and  $V_{max} = 0.77$  U / mg (Appendix 3).<sup>288</sup>

The resulting kinetic data attested of the His-HLADH and TbADH-WT integrity. Therefore, the results were set as a reference for further inhibition tests. TbADH-5M was not tested because of the absence of the catalytic  $Zn^{2+}$ .

### 5.3.2 *Investigation of the binding affinity between the transition metal catalysts and the enzymes*

In order to achieve the objective to non-covalently insert a catalyst in ADHs, its binding affinity inside the cofactor pocket is an important parameter to assess. After the design and synthesis of three metal catalysts, the next step was to investigate the binding affinity of their anchors inside NAD(P)H pocket.

A ligand-protein binding affinity can be assessed by direct means using various methods. For example, label tests typically use fluorescent labels on the ligand to analyse the binding. However, this requires expensive synthesis which can be challenging for the binding and integrity of the catalyst. Spectroscopic methods (e.g. X-ray, NMR) are also interesting tools to examine details on the binding site, the ligand orientation, conformation and interactions. These methods are however time consuming.

In this work, two different methods were tested to investigate the binding affinity of the three metal catalysts inside the ADHs. Firstly, the label free ITC experiment was used to directly measure the binding affinity of a ligand for a protein. Secondly an indirect binding affinity assessment was conducted by inhibition tests against NAD(P)<sup>+</sup>, the natural cofactor of ADHs.

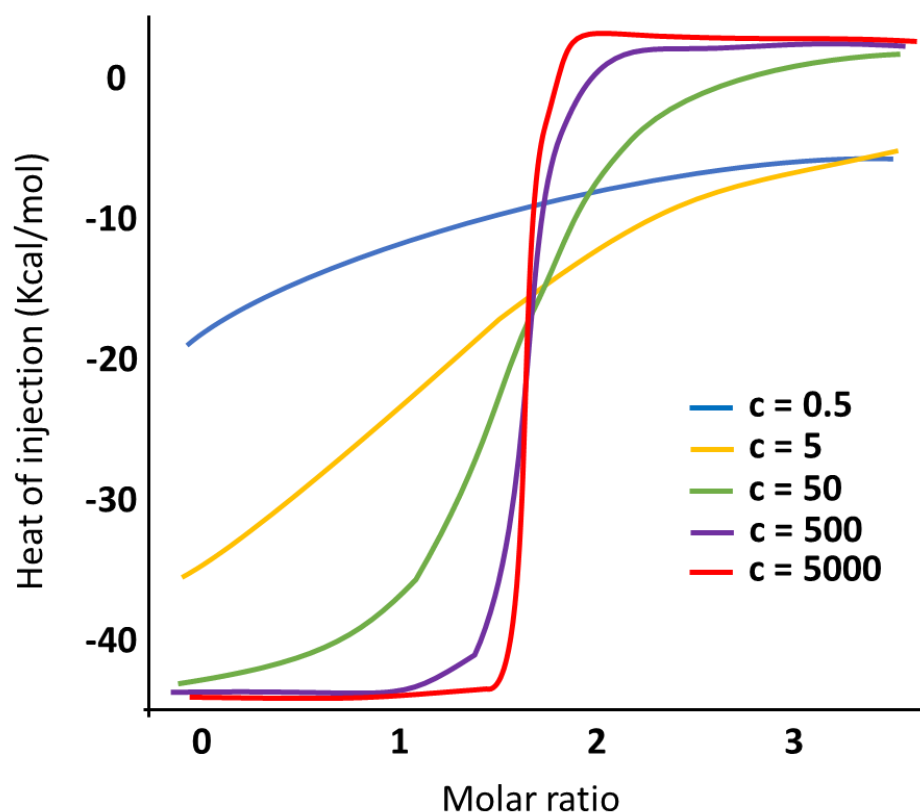
#### 5.3.2.1 Determination of the binding affinity of natural and synthesised cofactors with Isothermal Titration Calorimetry

ITC was used to assess the binding affinities of the catalyst anchoring part inside the enzymes scaffold. This method is frequently used to analyse the thermodynamic parameters of a binding interaction based on the measurement of heat absorbed or released upon binding. At constant pressure, the heat equals the enthalpy of binding. With the ITC technique it is possible to directly measure a dissociation constant ( $K_d$ ), and thus the binding affinity of a ligand for a protein. One disadvantage is the uncertainty to determine if the displayed ligand affinity accounts only for the targeted binding and not for other unspecific binding.

The resulting data are represented by a sigmoidal binding curve, where the slope corresponds to the association constant ( $1/K_d$ ) and the height to the enthalpy of binding. The binding curve and the dissociation constant depend on the concentration of both enzyme and ligands (**Figure 75**). A theoretical “c value” has been defined by the ratio of ligands concentration to the dissociation constant (equation (4), with  $n$  = stoichiometry):<sup>296</sup>

$$C = n[\text{titrand}] / K_d \quad (4)$$





**Figure 75:** Representation of the  $c$ -value influence on the sigmoidal curve. When the slope is too shallow ( $c = 0.5$ ) or too steep ( $c = 5000$ ), the  $K_d$  can't be defined with accuracy.<sup>297</sup>

From literature, the determination of  $K_d$  becomes reliable when the  $c$ -window is higher than 40 and under 1000, with a sigmoidal curve.

The first set of experiments was a blank control with no ligands in the dropping from the syringe solution. These injections showed the artefacts of dilution from the syringe content in the cell solution. The blank was then deduced from the experiments. In the second set of experiments, both forms of natural cofactors were investigated to set a strong binding reference with the three ADHs.

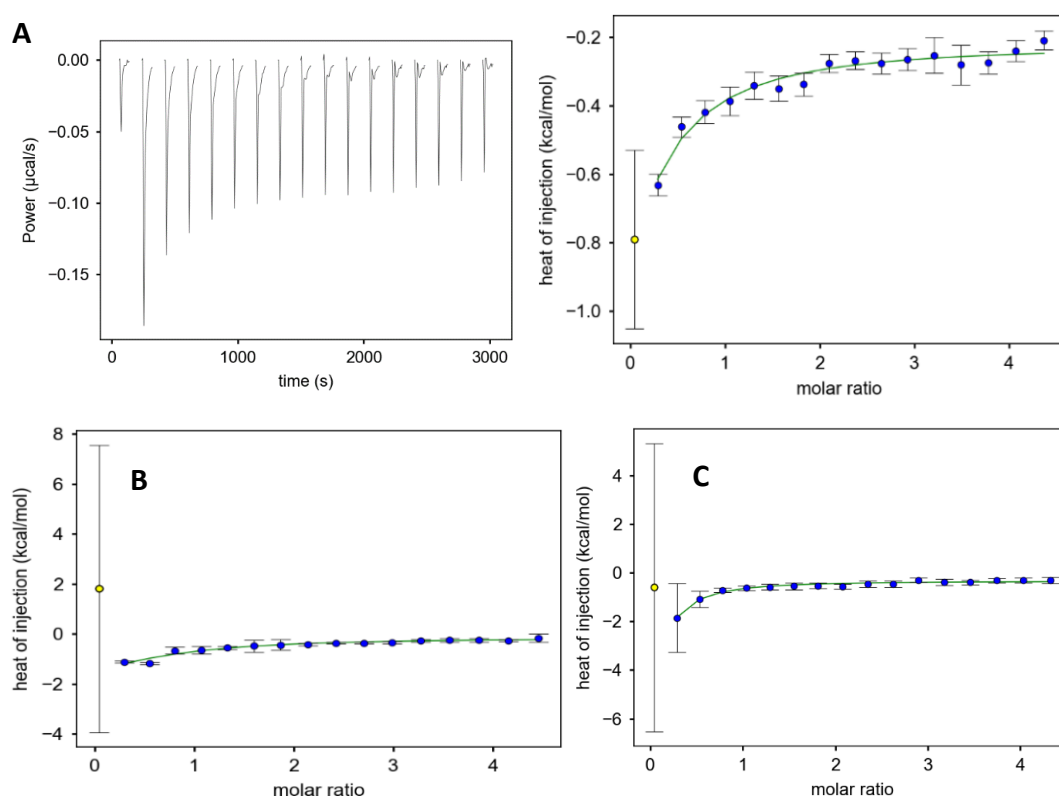
The final set of experiments measured the affinity of the previously synthesised deprotected carboxylic acid anchor **36** and the Boc protected  $\beta$ -diketoester anchor **38**. Compound **38** was not deprotected from the Boc group, but the anchoring part  $\beta$ -diketoester was interesting to test for affinity.

#### - *ITC of the three enzymes with their natural cofactor*

No data were found in literature for TbADH-WT ITC experiment testing the affinity of its natural cofactor NADP(H). The NADPH binding was previously assessed by ITC with an ADH yeast with an enzyme concentration of 50  $\mu\text{M}$ .<sup>93</sup> Only  $\text{NAD}^+$  binding was reported for His-HLADH, with an enzyme concentration of 150  $\mu\text{M}$ , by Ross and co-

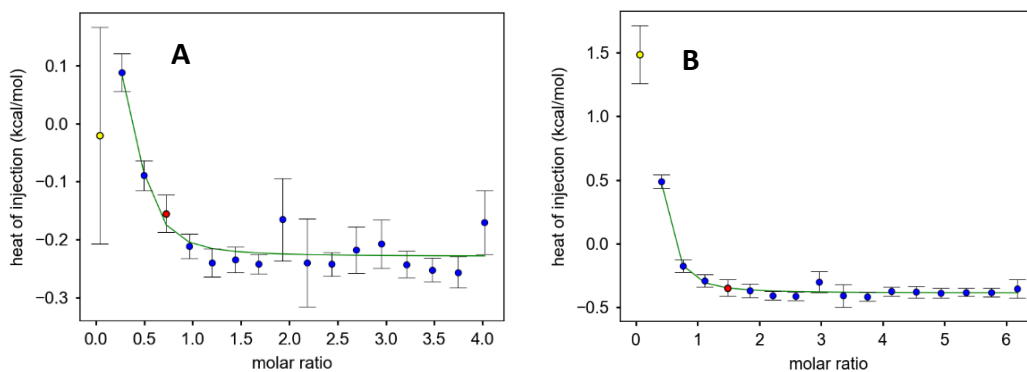
workers.<sup>298,299</sup> In a standard ITC experiment, the standard concentrations of enzymes are 10-30  $\mu\text{M}$  with a 10-fold concentration of ligand. From these data, a starting concentration of 50  $\mu\text{M}$  enzymes was attempted, with no relevant signal for both His-HLADH and TbADH-WT.

At the level of 100  $\mu\text{M}$ , the natural NAD(P)H cofactors start to show an affinity for all the enzymes (**Figure 76**). However, without a better binding curve shape, it was impossible to determine with precision a significant  $K_d$ .



**Figure 76:** ITC experiment with the enzyme natural cofactor at  $[E]=100\ \mu\text{M}$  and  $[\text{NAD(P)H}] = 2100\ \mu\text{M}$ . **A)** In TbADH-5M, the left panel displays an example of raw data of the heat pulses resulting from each titration. On the right, the curve represents the integration of the heat, normalised per mole of NADPH; **B)** TbADH-WT and **C)** His-HLADH. The molar ratio refers to titrant (ligands) / titrand (enzymes).

These results might imply a low binding affinity of the ADHs for their natural cofactor. Higher concentration of ligand and enzymes were necessary to improve the signal. Similar low affinities were obtained with the oxidised  $\text{NAD(P)}^+$  cofactors, which data showed a nearly flat curve (**Figure 77**).



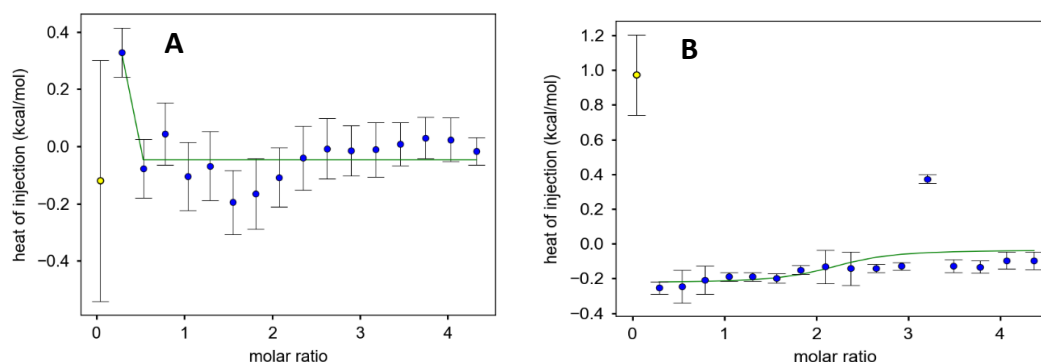
**Figure 77:** ITC experiment using 2100  $\mu\text{M}$   $\text{NADP}^+$  and 100  $\mu\text{M}$  enzyme concentration. **A)** TbADH-5M and **B)** TbADH-WT. The molar ratio refers to titrant (ligands) / titrand (enzymes).

The beginning of the curve was decreasing. With low significance, it was difficult to interpret the results as the first peak could refer to an artefact. A second possibility would be a change in conformation of the enzyme upon binding of  $\text{NADP}^+$ , impacting the binding of more ligand and thus reducing the heat. The conformational changes induced by  $\text{NAD}^+$  binding in HLADH was also hypothesised by Ross and co-worker<sup>298</sup> during calorimetry measurement. A change in conformation could enhance the affinity toward the cofactor, implying a higher affinity for  $\text{NADP}^+$  than  $\text{NADPH}$  or other ligands.

#### - *ITC of the three enzymes with the synthetic cofactors*

The synthesised anchors were tested in parallel to the natural cofactor.

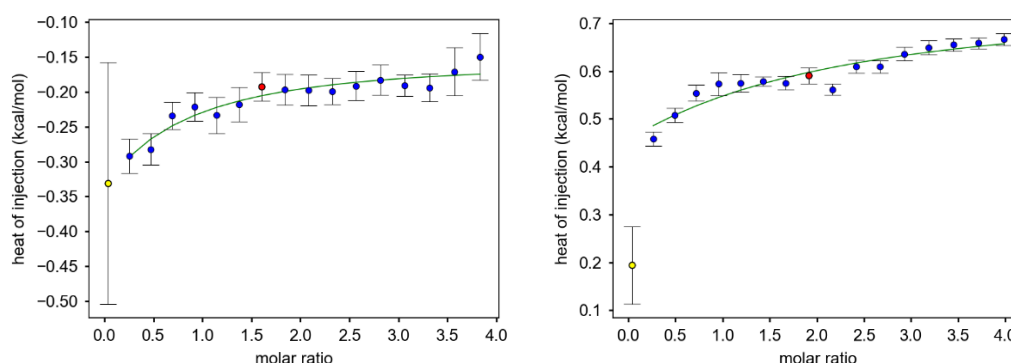
Similar to the natural cofactors test, the first attempts at low enzyme concentration gave inconclusive results for both products **36** and **38** (flat line). An increase in concentrations at  $[\text{E}] = 100 \mu\text{M}$  and  $[\text{L}] = 2100 \mu\text{M}$  also provided a flat binding curve (**Figure 78**), whereas experiments with  $\text{NADPH}$  started to show a slight curve. These observations were in agreement with previous docking results suggesting the designed anchors bound with weaker affinity to the enzymes than  $\text{NAD(P)H}$ .



**Figure 78:** Example of ITC experiments at  $[E] = 100 \mu\text{M}$  and  $[L] = 2100 \mu\text{M}$  with **A)**  $E = \text{His-HLADH}$ ,  $L = \text{acid } 36$  and **B)**  $E = \text{TbADH-WT}$ ,  $L = \beta\text{-diketoacid } 38$ . The molar ratio refers to titrant (ligands) / titrand (enzymes).

The enzyme concentration was again increased to  $130 \mu\text{M}$  with a ligand concentration of  $2400 \mu\text{M}$ . For the acid **36**, a flat line was again observed with both TbADH-WT and His-HLADH.

The experiments with anchor **38** started to show a shallow slope (**Figure 79**). This proved a start of binding affinity, even if the curve was less steep than the experiment with the natural cofactors.



**Figure 79:** ITC experiment with  $[E] = 130 \mu\text{M}$  and  $[L] = 2400 \mu\text{M}$  with **A)**  $E = \text{TbADH-WT}$ ,  $L = \beta\text{-diketoacid } 38$  **B)**  $E = \text{His-HLADH}$ ,  $L = \beta\text{-diketoacid } 38$ . The molar ratio refers to titrant (ligands) / titrand (enzymes).

The concentrations of enzymes applied in the experiments were again not high enough for a relevant binding curve slope. Thus, the  $K_d$  and thermodynamics reported contained significant uncertainty.

From the data, cautious conclusions can be attempted. First, with absence of signal, the benzoic acid **36** had weak or no affinity for the enzymes. With a shallow slope, the anchor **38** seemed to bind with better affinity to the enzymes than the shorter benzoic acid anchor **36**. Nonetheless, both bound to the enzymes with a weaker affinity compared to the natural NAD(P)H cofactor. These conclusions were also in

accord with the *in silico* study findings, where NAD(P)H showed the highest affinity score and the ligands with several hydrophilic groups (hydroxyls and carboxylic acids) were ranked just after.

The ITC binding curve quality can mainly be enhanced by increasing the amount of protein (to 150-200  $\mu$ M) and ligands. Encouraging reported data with 150  $\mu$ M or 10 mg/mL concentration of other ADHs types strengthen this hypothesis.<sup>299,300</sup> However, with increase in concentration, the proteins might aggregate in solution and thus reduce the binding capacity. An investigation of the enzyme behaviour would thus give input on the experiment quality. The high amount of enzymes and ligands needed for this experiment conducted to no further ITC assays.

#### 5.3.2.2 Indirect determination of the binding affinity of synthesised compounds to ADHs, by inhibitory assays

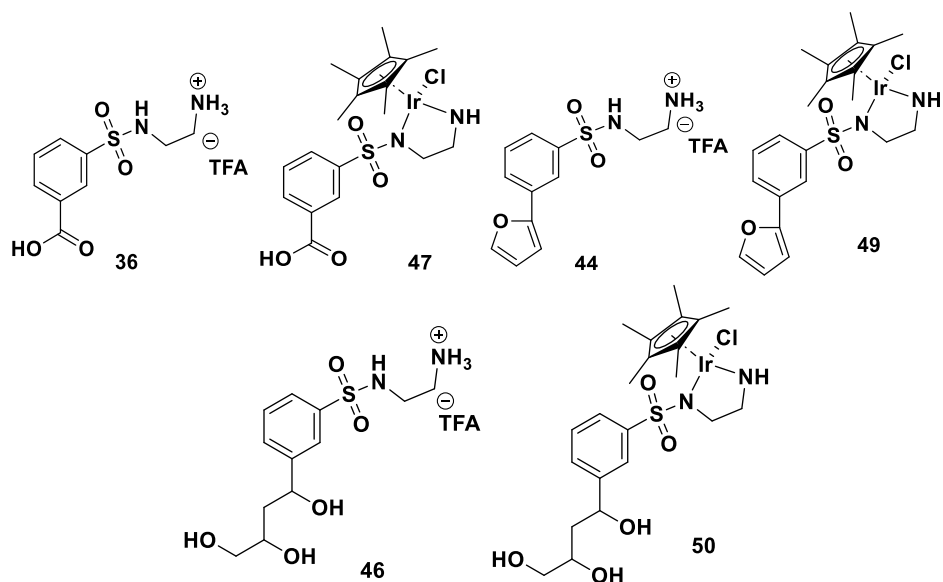
With inconclusive ITC results, the catalysts access and affinity to the cofactor binding pocket was investigated by indirect means using inhibition assays.

Assuming the synthesised ligands act as competitive inhibitors, their binding to a protein active site will prevent the natural reaction to occur. They would replace NAD(P)<sup>+</sup> by binding in its pocket, but without its reactivity due to the loss of the nicotinamide function.

The inhibitory potential of a ligand gives an input in its binding capacity with small amounts of enzymes and ligands. However, this method of affinity measurement is indirect and depends on the affinity of the enzyme for its natural cofactor.

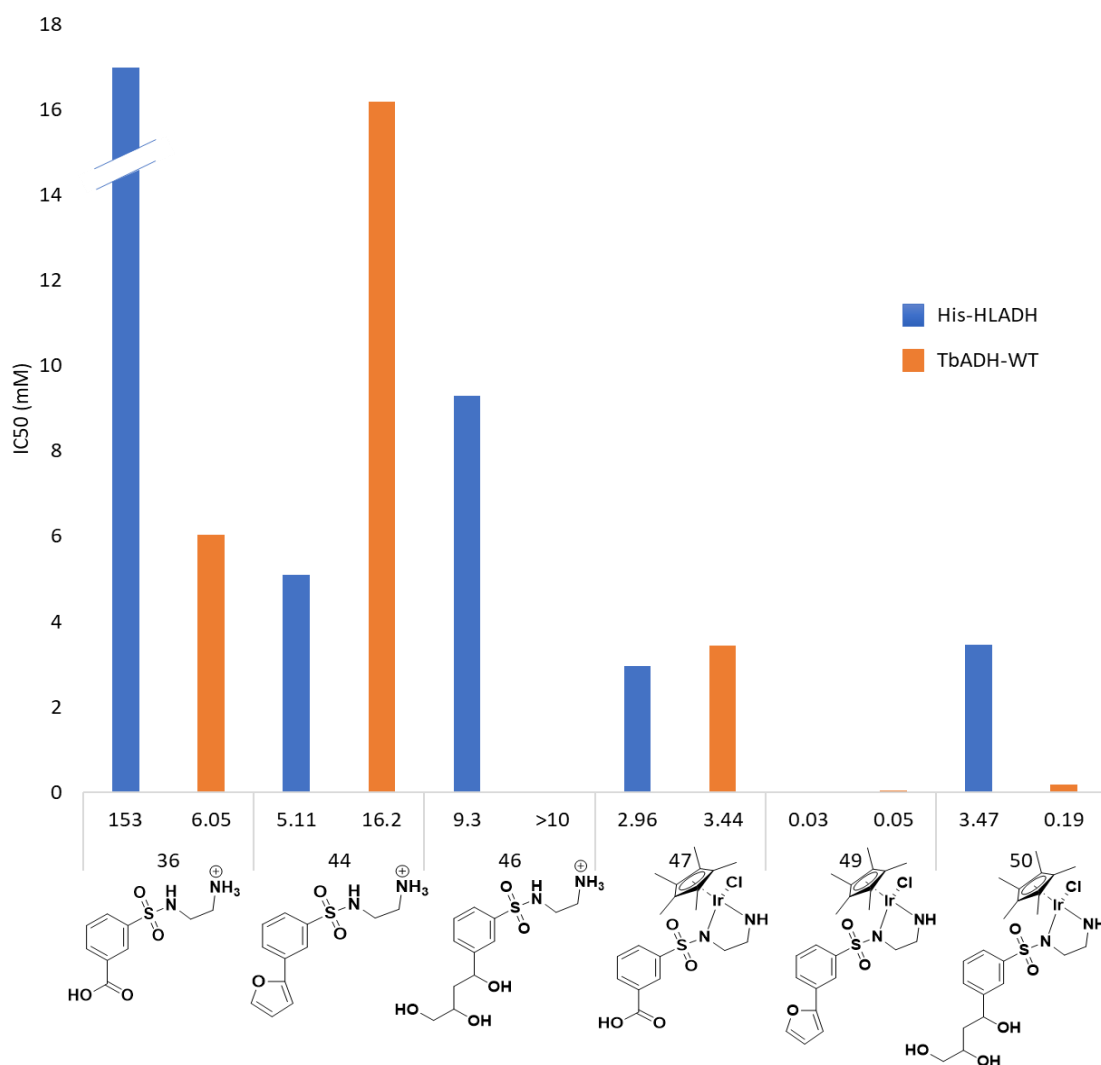
Inhibition of both His-HLADH and TbADH-WT activity was tested in the oxidative reaction of ethanol and butan-2-ol respectively. Without Zn<sup>2+</sup>, TbADH-5M was unfunctional and not tested in these experiments.

The diamine anchors (compounds **36**, **44**, **46**, **Figure 80**) were designed through docking to anchor the metal with high affinity, catalytic orientation and location (Chapter 3, page 57). Therefore, their affinity was tested alongside the transition metal complexes (compounds **47**, **49**, **50**, **Figure 80**). The large size of the iridium functionality is likely to have an impact on the binding inside the tight cofactor binding pocket.



**Figure 80:** Synthesised anchors and their corresponding iridium complexes used in this affinity study. Both types of compounds will be tested for their supramolecular binding strength and specificity within the wild type TbADH-WT and His-HLADH.

Several concentrations of each ligand were evaluated against a constant concentration of cofactor (0.5 mM) in the presence of the substrate. The formation of NADH was followed at 340 nm and the data were plotted on a logarithmic scale (graphs in Appendix 4). A sigmoidal fit was computed with the Prism software and the  $IC_{50}$  of each inhibitor was calculated for the two enzymes with a good parameter fit (SD replicates between 0.01 and 0.1, **Figure 81**). The  $IC_{50}$  corresponds to the concentration of inhibitor required to inhibit half of the enzyme activity, in competition with  $NAD(P)^+$ .



**Figure 81:** IC<sub>50</sub> values of synthesised anchors and iridium complexes with His-HLADH (blue) and TbADH-WT (orange) in the oxidative reaction. NAD(P)<sup>+</sup> (0.5 mM), ethanol (4 mM) and butan-2-ol (150 mM) were used with His-HLADH (15 µg) and TbADH-WT (1 µg) in 100 mM TrisHCl pH 8 buffer. The molecules ranged from 10<sup>-4</sup> to 10<sup>2</sup> mM (> 10mM for compound **46**: not enough stocks to test further).

Deprotected diamine anchors (**36**, **44** and **46**) expressed an IC<sub>50</sub> between 5 mM and more than 10 mM for both His-HLADH and TbADH-WT. These results indicate a low inhibition to the enzymes and so a low affinity in comparison with the natural NAD(P)<sup>+</sup> cofactor. The corresponding iridium complexes (**47**, **49** and **50**) showed higher inhibitions between 0.03 mM and 3.5 mM. The furan anchor with or without iridium metal (**44** and **49**) is the best inhibitor for both TbADH-WT and His-HLADH. The Triol anchor (**46** and **50**) follows, finishing by the acid (**36** and **47**).

The first scenario that could explain these observed inhibitions was the binding of the metal complex anchoring part to ADHs. This could have occurred either in the cofactor pocket or *via* unspecific inhibiting binding to other sites.

For example, in the case of catalyst **49**, the lowest IC<sub>50</sub> could be explained by the furan hydrophobicity favouring **49** binding in the substrate pocket, mainly hydrophobic.<sup>217</sup> As anticipated by the docking study, any hydrophobic part of a ligand tends to be introduced inside the substrate pocket, inducing an inhibition by lack of space for the substrate to bind (section 3.3.1, page 63). Thus, **49** higher inhibitions would come from competition with the substrate, which is less specific than competing with the natural cofactor. By comparison, the more hydrophilic compounds **50** and **47** could bind properly inside the cofactor pocket with a stronger competition against the natural cofactor, explaining a lower inhibition.

This binding of the anchor hypothesis can be challenged by the differences in inhibition between the deprotected and the metalated compounds. With a higher affinity observed when the iridium was present, it can be speculated a second scenario of an impact of the metal on the binding. d<sup>6</sup> piano-stool metals iridium and rhodium are known to bind to enzymes.<sup>301</sup> If the binding occurs in an inhibiting position (i.e. entrance of a pocket) the NAD(P)<sup>+</sup> or the substrate might be excluded from the pocket and not able to react.

The second scenario would be an aggregation promoted by the iridium that would inhibit the enzyme activity. However, no precipitation in any of the iridium complexes solutions discarded this last hypothesis.

In conclusion, the iridium catalysts **49** and **50** inhibited the two wild type enzymes. To explain this inhibition, two hypotheses could be established:

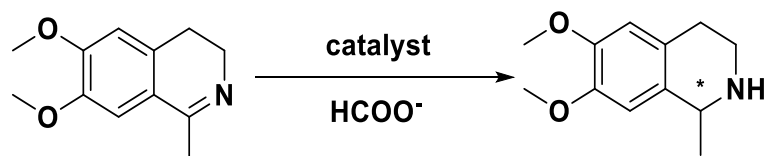
- the anchoring part of the catalyst bound in the cofactor or substrate pocket, preventing their binding and the natural reaction.
- the catalysts bound by the iridium in a position inhibiting the reaction.

Certainty on how the binding happen and the affinity for the ADHs cannot be defined because of the indirect analysis. More rigorous experiments, like crystallography, could further be attempted for details on the location and orientation of the bound ligands.



### 5.3.3 Assessing the catalytic reactivity of the three metal catalysts and the AMs

The inhibition observed in the previous section by the metal catalysts could reflect an affinity for the wild type ADHs. The reactivity of the AMs will here be tested by catalytic assays for cyclic imine reduction (**Scheme 23**).



**catalyst= AMs, 47, 49 or 50**

**Scheme 23:** Transfer hydrogenation of 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline to produce the chiral amines 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroquinoline, catalysed by the artificial metalloenzymes or the iridium catalysts alone, adapted from a previously describe protocol by Ward and co-workers.<sup>291</sup>

Before starting the catalytic experiments, a standard curve was established to analyse the HPLC results quantitatively (Appendix 5). The substrate imine 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline and the two chiral amines products 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline were used and extracted following the protocol described in the catalytic assays method (section 5.2.2.5, page 166).

The imine retention time was at 11.18 minutes and the two amines retention times were at 19.59 minutes and 23.93 minutes.

#### 5.3.3.1 Determination of the catalytic activity of the synthesised iridium catalysts

HPLC kinetic assays for the synthesised metal complexes were performed in aqueous solution to evaluate their catalytic potential for cyclic imine reduction.

The three iridium catalysts (**47**, **49**, **50**) were tested for the catalytic transformation of the 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline into the racemic mixture of the amine 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline for 24h in aqueous media (curves in Appendix 5).

**Table 19:** Turnover frequency (TOF) of the three iridium catalysts is presented for the reduction of 4mM cyclic imine substrate with 50  $\mu$ M of catalysts at 38°C after 1h of reaction. The experiments were run in triplicates for 24h and the results presented are the mean  $\pm$  standard error.

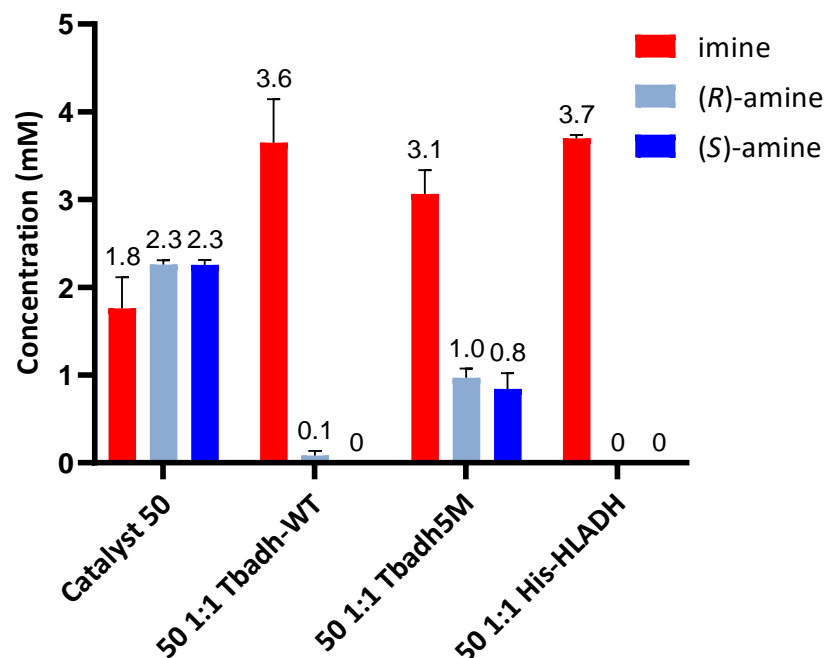
Catalysts	yield %	TOF ( $\text{min}^{-1}$ )	TON
<b>47</b>	54 $\pm$ 0.03	1.4 $\pm$ 0.11	59 $\pm$ 8.39
<b>49</b>	57 $\pm$ 0.04	1.8 $\pm$ 0.07	66 $\pm$ 1.45
<b>50</b>	94 $\pm$ 0.01	4.5 $\pm$ 0.09	90 $\pm$ 1.04

The reactivity of the three catalysts was validated for the aromatic imine reduction with no chiral induction. According to the results, the TOF depends on the anchor structure linked to the metal. The iridium catalyst **50** with a triol group was highly reactive compared to the iridium complex **47** with a benzoic acid anchoring part.

These activity data were satisfying compared to previous work in literature. Ward and co-workers convert the same isoquinoline substrate in racemic mixtures with 10 TON using a Cp\* iridium complex<sup>9</sup> and with 35 TON using a biotinylated iridium complex.<sup>64</sup> Outside the scope of this project, these synthesised iridium catalysts were of interest for the chemical reduction of imine to chiral amine building blocks, in aqueous media.

#### 5.3.3.2 Determination of the catalytic activity and specificity of assembled AMs.

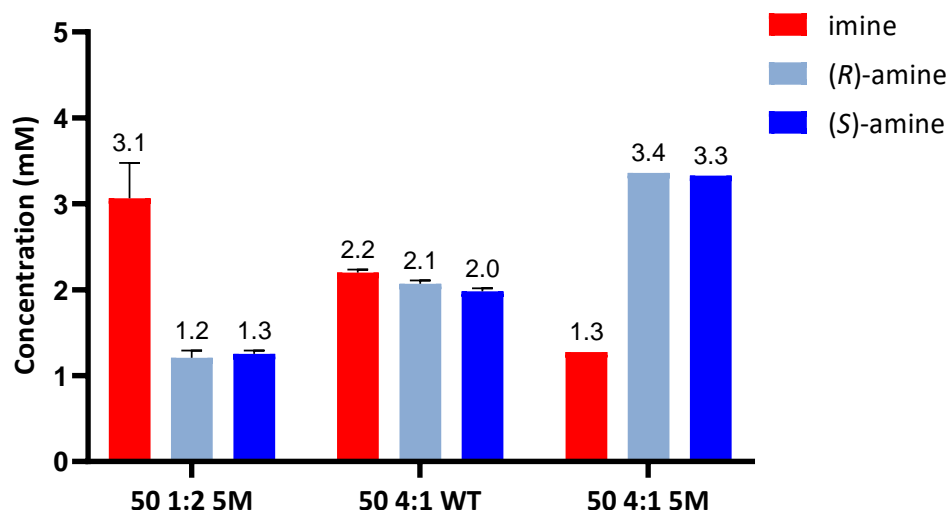
Investigation into the catalytic reactivity of the iridium transition metal complexes in combination with the three enzymes were performed. The protocol applied in the previous section was only modified in temperature with His-HLADH, which is more efficient at 25 °C.<sup>288</sup> The following graphs are presenting the three catalysts **47**, **49** and **50** in combination with TbADH-WT, His-HLADH and TbADH-5M at different concentrations.



**Figure 82:** The product and substrate concentrations obtained from the reduction of 4 mM cyclic imine at 38 °C in PBS buffer, catalysed by 50  $\mu$ M iridium complex **50** and 50  $\mu$ M enzymes, for 24h (25°C for His-HLADH reaction). The graph presents the iridium complex alone and in combination with the three selected enzymes for AMs engineering: His-HLADH, TbADH-WT and TbADH-5M.

This first graph displays the catalytic activity of iridium complex **50** alone and in a one-to-one concentration combination with the three ADHs (**Figure 82**). The catalyst alone presented a catalytic activity with no stereoselectivity like previously mentioned. No reaction was achieved with His-HLADH and only a very small quantity of amine was detected with TbADH-WT. Finally, formation of the amine was observed with TbADH-5M. No significant stereoselectivity was observed for all setups.

The activity displayed by TbADH-5M gave a first indication of success for the rational design of this AM. To further perturb the equilibrium towards the bound state inside the enzyme binding pocket, the concentration of enzyme was doubled. In parallel, a rise in the iridium complex concentration was also applied to saturate potential binding sites and verify the AM selectivity and possible inhibition (**Figure 83**).



**Figure 83:** The product and substrate concentrations obtained from the reduction of 4 mM cyclic imine at 38 °C in PBS buffer, for 24h. The graph presents the iridium complex **50** with TbADH-5M and TbADH-WT, at various concentrations combinations.

For an increased enzyme concentration, the catalysis was slowed down. It can be suggested that no free metal catalyst remained in the mixture and the observed activity should come from bounded catalyst.

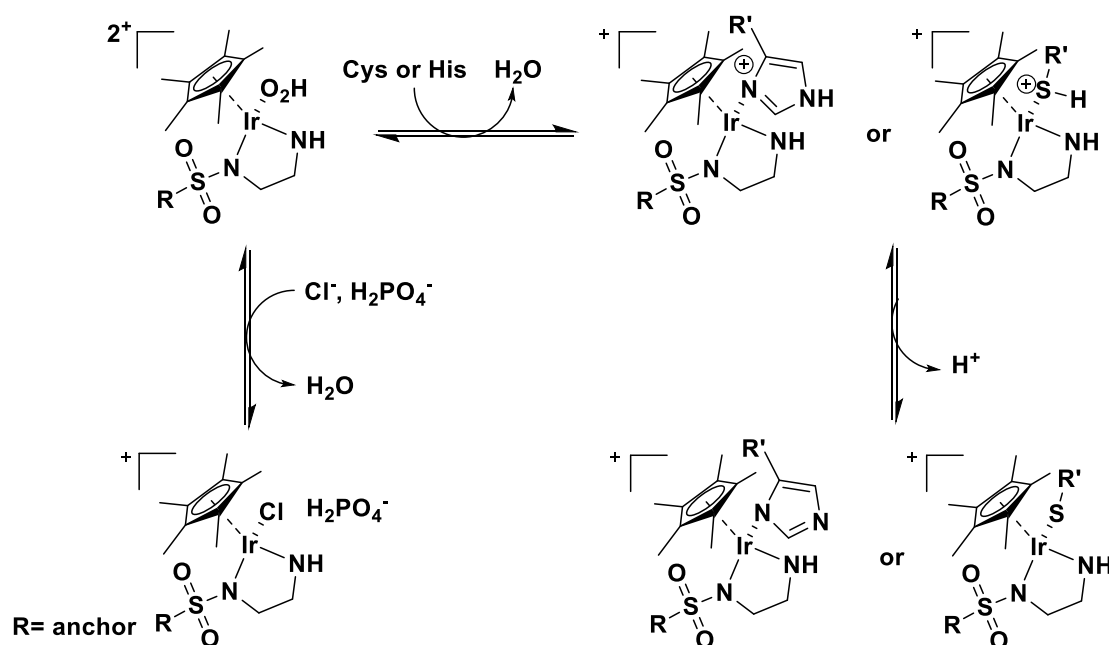
On the other hand, the production of amine was higher when the metal catalyst was four times the enzyme concentration. The wild-type enzyme taken as a reference showed activity this time, potentially indicating a reaction of the free metal only. In comparison TbADH-5M got higher conversion, indicating an activity from the free and bounded catalyst in excess. This suggested again that the reaction proceeded with metal catalyst bound to TbADH-5M, even though with lower turnover, whereas in the wild type, the reaction was mainly inhibited by the enzyme.

From these results, two main observations could be made: in combination with wild type enzymes, the catalytic activity of the transition metal catalyst was inhibited. On the other hand, the mutant TbADH-5M proved to form a functional AM with reproducibility but no stereoselectivity.

The absence of activity with the wild type enzymes correlated with the binding inhibition observed in the preceding inhibition tests. This lack of activity can be explained by different scenarios already considered previously (section 5.3.2.2, page 176). The first option is a binding of the catalyst **50** to the enzyme through its metal iridium, with a mutual inhibition. The second scenario is the expected binding of **50**

anchoring part inside the cofactor pocket, but without space for the substrate to approach the catalyst.

In the first hypothesis, the iridium metal could have reacted with amino acids in the enzyme (mainly cysteine and histidine), blocking the catalytic activity. The chloride linked to the iridium is known to be hydrolysed in aqueous media.<sup>301</sup> This could represent an activation step, followed by the reactivity of the iridium to amino acids. Hollmann and co-workers developed a theoretical mechanism for the mutual inhibition of rhodium catalyst and an ADH.<sup>301</sup> This mechanism can be extended to similar iridium complexes, which have also been implicated in monoamine oxidases inhibition.<sup>70</sup> The theory implies the hydrolysis of chloride, followed by metal coordination with cysteine or histidine residues which induce the inhibition of the metal catalytic activity (**Figure 84**). In addition, if these amino acids were involved in catalysis or positioned in a cofactor / substrate pocket, this would also provoke the protein inhibition.



**Figure 84:** Proposed mechanism of the inhibition of  $d^6$  piano-stool iridium complexes by binding to amino acids, adapted from Hollmann and co-workers theory.<sup>301</sup>

In this project, this hypothesis was supported by the observed activity of catalyst **50** in the five-cysteine mutant TbADH-5M, when no activity was found in the wild types. If no cysteine were present, the iridium catalyst was free to react with the imine substrate without binding to an amino acid.

In the same way, His-HLADH AM being totally inactive, the binding of iridium to the polyhistidine tag could add to this theory. The His-tag is well known to bind to cation transition metal ion.<sup>302</sup> Many researches showed examples of metalloenzymes and zinc, copper, nickel metals being coordinated by the His-tag.<sup>303,304,305,306</sup> In this project, the natural  $\text{Zn}^{2+}$  was not affected by the His-tag (proved by kinetic assays) . However, the introduction of iridium with weakly bound ligands could have followed the previously described mechanism and bind to histidine (**Figure 84**).

Cleavage of the tag to obtain a closer to nature HLADH or the use of commercially available wild type HLADH would discard any potential impact of the His-tag. Another solution would be the use of acidic conditions (pH= 5) to avoid the metal binding. Those conditions might however affect the enzyme, which activity is optimal at pH = 8.8.<sup>288</sup>

The metal binding theory is however also debatable. Other authors have succeeded in using these transition metal complexes with TbADH and HLADH without any mutual inhibition.<sup>301,307</sup> Besides, Hollmann and co-workers used a four-time ratio of metal complex to obtain inhibition, where in this experiment, the stoichiometry was equivalent and no inhibition was observed at a four to one ratio.

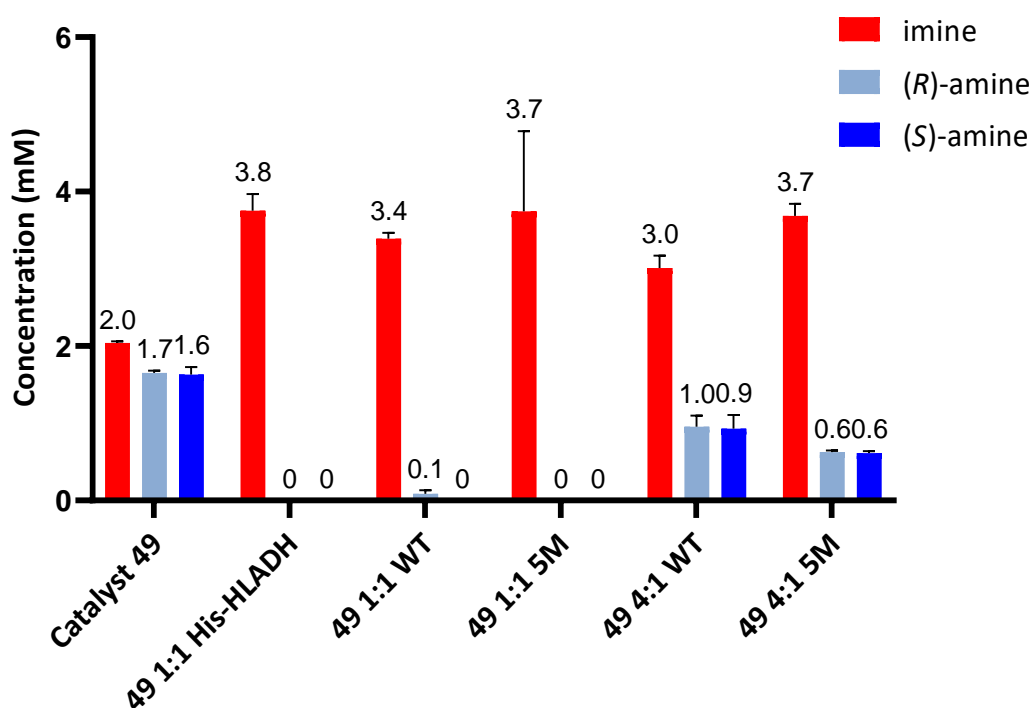
Thus, the second plausible theory could be a binding of the anchoring part of **50**, without possible transfer hydrogenation reaction being performed in the wild types. For example, mutual inhibition could come from a binding of the catalyst in the substrate pocket, in an unspecific allosteric site or in the cofactor pocket, but with obstacles to meet the substrate. This final proposition was endorsed by the activity showed with the mutant TbADH-5M as scaffold. The only structural difference with the TbADH-WT is the lack of catalytic  $\text{Zn}^{2+}$  and the space made in the catalytic area by mutation of H59 and D150 into alanine residues. As already suspected in docking and in the choice of mutated enzymes, more space inside the catalytic area allowed meeting of the substrate and metal catalyst, resulting in amine production (i.e. **Figure 72**, page 160).

However, precautions need to be taken as the reaction with TbADH-5M showed no stereoselectivity, suggesting the possibility of a reaction catalysed by free catalysts

outside the enzyme. In addition, the integrity of TbADH-5M activity was not tested prior to catalysis experiments due to the lack of catalytic  $\text{Zn}^{2+}$ .

To properly investigate the mutual inhibition, catalytic activity could further be assessed with amino acid (particularly cysteine) in the mixture. If the activity decreases, binding with cysteine should have occurred, which verify the first hypothesis of inhibition by iridium binding. Besides, adding NADH as a competitor in the case of TbADH-5M catalysis experiment could affect the reactivity if the binding occurs inside the pocket.

The two other metal catalysts (**47** and **49**) were also tested for their reactivity within the three ADHs, following the same protocol (**Figure 85** and **Figure 86**).



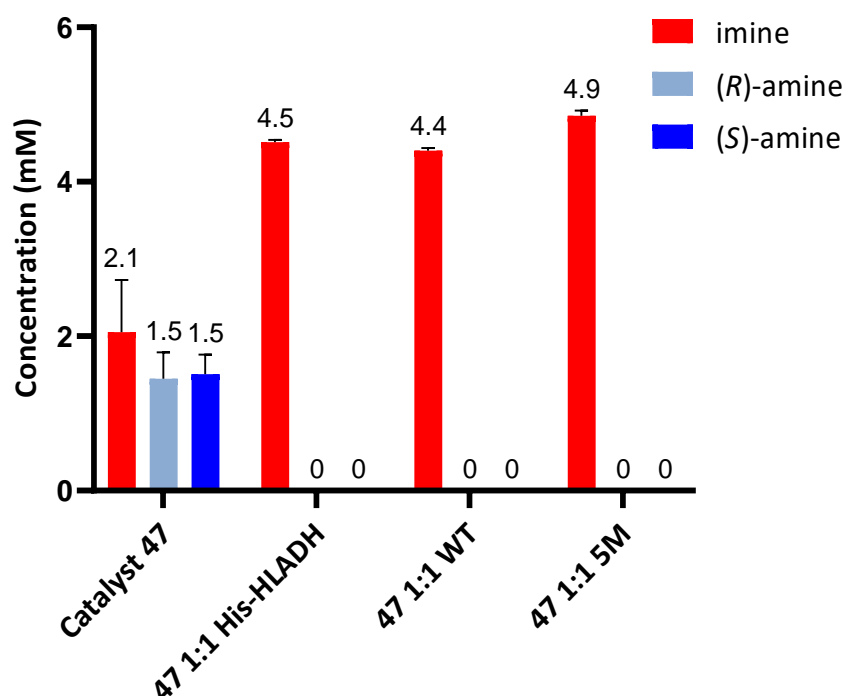
**Figure 85:** The product and substrate concentrations obtained from the reduction of 4 mM cyclic imine at 38 °C in PBS buffer, catalysed by iridium complex **49** alone and in combination with the three enzymes, for 24h (25°C for His-HLADH reaction).

Catalyst **49** complex alone was active with no stereoselectivity. In combination with all the three enzymes, the iridium complex **49** gave low to no activity. When the concentrations of metal catalyst **49** were raised, again low activities were shown, especially compared to the same experiment with catalyst **50**, which obtained a complete conversion. Thus, it can be suggested that the observed small activity of **49** catalysis came from free metal catalyst in solution. These inhibition results in

combination with IC<sub>50</sub> data highlighted a binding of **49** with all the ADHs and a mutual inhibition.

With no activity showed by TbADH-5M lacking cysteines residues, the theory of a mutual inhibition by iridium binding was again less plausible. Thus, the second scenario of a binding in a non-catalytic way was again privileged. Contrary to catalyst **50**, which may bind properly in the cofactor pocket, the lack of activity of **49** with TbADH-5M suggest another inhibiting binding. Prior conclusions from inhibition assays tended to hypothesise the binding of **49** inside the substrate pocket, in correlation with docking studies. Compound **49** being hydrophobic, it would take the substrate place and thus cannot realise the imine reduction while inhibiting the ADHs natural activity.

Finally, catalytic tests on compound **47** also presented no activity in combination with all the enzymes (Figure 86).



**Figure 86:** The product and substrate concentrations obtained from the reduction of 4 mM imine at 38 °C in PBS buffer, catalysed by 50 μM iridium complex **47** alone and with 50 μM of the three enzymes His-HLADH, TbADH-WT and TbADH-5M, for 24h (25°C for His-HLADH reaction).

The total lack of catalytic activity, compared to the complex **50**, again highlight a difference in the binding mode to ADHs. The inhibition of the catalyst **47**, even without Zn<sup>2+</sup>, tends toward similar binding hypothesis than catalyst **49**: an unspecific



binding of the anchor part. As this catalyst represented a truncated intermediate structure, with no designed features for the binding in NAD(P)H pocket, the correct binding might not be supported, which led to unspecific inhibitory binding instead (e.g. in the substrate pocket). These data were in correlation with docking studies where **47** was part of the designed small library of catalysts but was not selected in the lead structures because of a very low ranking in affinity score.

## 5.4 Conclusion

The three selected AMs (TbADH-WT, TbADH-5M and His-HLADH) were successfully expressed and purified. The two wild type enzymes integrity and functionality were validated through kinetic parameters consistent with literature.

Two assays were run to assess the binding affinity of the synthesised catalysts for the three ADHs. A direct measurement with ITC assay was not able to validate binding affinity. The second indirect experiment in form of inhibition assays of the catalysts against the catalytic activity of the natural cofactor NAD(P)<sup>+</sup> showed binding affinities in the low mM regime for the metal catalysts (**47**, **49** and **50**) with the two wild type enzymes. Poorer binding affinity was displayed by the shorter ligands **36** and **47**, which was consistent with the *in silico* study and previous work from M. Basle.<sup>109</sup>

Finally, the catalytic activity of the three metal catalysts (**47**, **49** and **50**) and the resulting AMs was assessed by HPLC monitoring of the products of a catalysed imine reduction in aqueous media. The three iridium catalysts alone were active for the imine reduction with no stereoselectivity.

From the catalytic data of the conjugated metal with ADHs, a more complicated picture arose.

No activity was found for all catalysts in combination with His-HLADH and TbADH-WT. A clear but lower activity was obtained by the catalyst **50** in TbADH-5M, with no stereoselectivity. Further experiments with higher concentrations of the catalyst **50** and comparison with the other metal catalysts and inhibition studies confirmed that the activity was unlikely to come from free catalyst in solution. Thus, the catalyst **50** was deemed to have bound to the enzyme in an active way, revealing the design of an AM.

The lack of stereoselectivity for the TbADH-5M – **50** AM could be explained by an active binding of the catalyst **50** to TbADH-5M, but not in the deemed cofactor binding pocket. For example, a surface exposed binding could explain the lack of stereoselectivity of the designed AM. However, it remained unclear where this binding site would be located. It is also not directly supported by the inhibition studies that clearly show direct inhibition, suggesting binding in the cofactor or substrate binding pocket.

The two other catalysts **47** and **49** lacked activity with all enzymes, suggesting a different inhibitory pathway, and again supporting the assumption of an active AM with catalyst **50**.

Three scenarios were developed to explain the inhibiting binding modes:

- The iridium metal bound to amino acids (e.g. cysteine) and provoked mutual inhibition of the catalyst and enzyme. This theory was invalidated by literature<sup>301,307</sup> and by the lack of activity from **47** and **49** with TbADH-5M, yet missing cysteines.
- The metal catalyst anchoring part made unspecific binding contacts in an inhibiting position, notably inside the hydrophobic substrate binding pocket. The truncated catalyst **47** and the hydrophobic catalyst **49** were perfect candidates for this hypothesis. This was supported by the combination of lack of activity for **47** and **49** based AM constructs with existing inhibition of the natural activity in both cases.
- The metal catalyst **50** binds properly inside the cofactor pocket, but the contacts with the substrate in wild type enzymes are prohibited by a lack of space due to  $\text{Zn}^{2+}$ . This last hypothesis was supported by the effective catalytic assays of **50**, only in TbADH-5M where  $\text{Zn}^{2+}$  was missing.

## 6 Discussion, conclusions and perspectives

This thesis presents the rational design of artificial metalloenzymes (AMs) for the transfer hydrogenation of imines. A supramolecular anchoring strategy was chosen for the introduction of a transition metal catalyst inside an enzyme scaffold based on alcohol dehydrogenases (ADHs). The reactivity was introduced by a d<sup>6</sup> piano-stool iridium metal complex.

The initial focus of this project was the creation of small anchors with high affinity for the NAD(P)H binding pocket of ADHs. The final objective was then to metalate these anchors to form imine reduction catalysts and non-covalently bind the resulting complexes inside ADHs scaffold. The engineered AM should be new, robust, and efficient toward the catalytic reduction of imine into chiral amines.

Anchors of the transition metal catalysts were first designed by computational modelling studies focused on the development of docking protocols. The Horse Liver ADH (HLADH) and the *Thermoanaerobacter brockii* ADH (TbADH) were chosen as scaffolds due to their large substrate scope and the similarity of their natural reaction (aldehyde reduction) to the targeted imine reduction. TbADH is also both thermostable and resistant to solvents,<sup>97</sup> which is advantageous in a synthesis pathway where high temperatures and organic solvents are commonplace. In addition, by process of evolution, the candidate enzymes have a natural high affinity to bind their native NAD(P)H cofactors, therefore the docking focused on finding high affinity NAD(P)H analogues for the supramolecular catalyst anchoring.

A docking protocol was developed and validated for the reliable approximation of affinity inside the ADHs cofactor binding pocket. The natural cofactor, online libraries and fragments of NAD(P)H were docked as a starting point of the *in silico* study. Enzymes scaffold with and without their catalytic Zn<sup>2+</sup> were considered. The space offered by lacking the zinc ion could indeed be a potential advantage to insert the metal catalyst further toward the substrate.

In addition, the presence of substrate in the co-crystallised structure was essential to the docking. The substrate pocket being essentially hydrophobic, any hydrophobic structures were inserted inside the pocket, obstructing the substrate insertion.

A model for anchors was designed based on conclusions from the first docking screening. The typical structure was composed of an aromatic ring substituted with a hydrophilic anchoring part on one side, and a sulfonamide support for the metal on the other side. Hydrophilic functional groups are necessary for maximising the interactions inside NAD(P)H binding pocket. This was supported by experimental results from the literature, suggesting low affinity and activity for truncated hydrophobic mimics of NADH inside HLADH.<sup>106,109</sup> Hydrophobic features were also discarded because of their potential insertion in the hydrophobic substrate pocket (inducing possible inhibition of substrate binding).

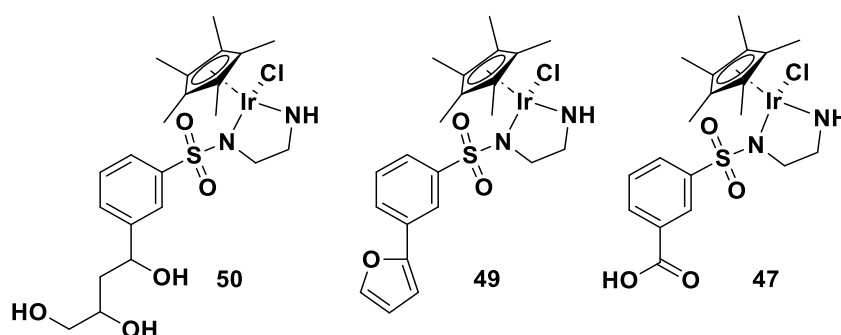
The resulting library of anchors was screened inside ADHs with a new docking protocol considering the affinity score, the location and orientation of the anchors. A catalytic position of the metal catalyst was indeed essential to the activity. This created docking protocol includes first the use of the Glide XP method for anchor affinity screening, then the Induce Fit Docking (IFD) method was applied to investigate the binding affinity and orientation of metalated anchors. The IFD method was necessary to introduce flexibility into the enzyme and insert the metal. However, effect induces by the binding of a metal on the electronic state and geometry of the first coordination sphere, the ligand and the scaffold are considerations introducing very high complexity for the molecular modelling of AMs.<sup>308</sup> Thus, the results have to be considered carefully as approximations with crucial information on anchor structure model and hydrogen bonding. Nevertheless, a set of lead anchor structures was selected based on their affinity score, the location and orientation of the metal.

In future, refinement with more rigorous experiments may be necessary, for example to improve the detailed knowledge of the catalytic reaction and the catalysts binding process. Molecular dynamic simulations or QM/MM are increasingly used methods for the in depth analysis of a binding site and catalytic reaction, however at the costs of higher time and computational resource consumption.<sup>165,308</sup> In addition, should an appropriate anchor be discovered experimentally, new computational experiments can also be considered for further optimisations. For example, genetic mutations of residues in the cofactor pocket or modification of anchor shape would likely increase binding capacities and stereoselectivity.

Based on the docking, metal complexes coordinated with the anchoring structures were synthesised. This objective included the design of a short divergent synthetic pathway able to provide a small library of anchors.

The key intermediate  $\beta$ -diketoester showed the disadvantage of its sensitivity to the applied and necessary deprotection conditions. The subsequent solution chosen was the reduction of this anchoring part to a ketodiol that is more resistant and still close to the initial anchor design. A complete modification of the key intermediate  $\beta$ -diketoester could also have been a solution. However, this would have varied entirely the anchoring part, possibly moving away from the docking models.

The subsequent metalation provided three catalysts: Ir-triol **50**, Ir-acid **47** and Ir-furan **49** (**Error! Reference source not found.**).



**Figure 87:** Synthesised metal catalyst complexes: Ir-triol **50**, Ir-acid **47** and Ir-furan **49**.

Based on previous docking conclusions, catalyst **50** met the requirement of several hydrophilic features and looked similar to the selected lead structures, missing only a carboxylic acid. Catalyst **49** and **47** were also interesting for a comparison in the binding affinity of selected hit structures and hydrophobic or truncated ligands.

Further possible chemical optimisations of the first coordination sphere are decisive, especially coupled with genetic optimisations, as observed in the work of many groups.<sup>49,58,63</sup> In addition, optimisations in the previous *in silico* study could also result in a new synthesis path toward different catalysts libraries.

The final part of this project focused on the construction of efficient AMs that aimed to reduce cyclic imine into chiral amines. Three ADHs available in the laboratory were expressed and evaluated as scaffolds: two wild types His-HLADH and TbADH-WT, and a mutant version of the wild type TbADH called TbADH-5M. The mutation resulted in the loss the  $\text{Zn}^{2+}$  ion, thus giving more space for the introduction of the iridium in the

catalytic site. Two main parameters were then analysed in this chapter: the binding affinity of the synthesised catalysts inside ADHs and the catalytic activity of the resulting AMs.

First, a direct and label-free means was run to test affinity by Isothermal Titration Calorimetry (ITC). This robust technique would provide the binding, kinetics and thermodynamic information needed to evaluate the anchoring part of the catalyst with high accuracy.<sup>296,309</sup> Although the assay was sensitive to detect binding affinity, inconclusive data were retrieved due to weak signals displayed by the binding enthalpy between the enzymes and NAD(P)H. This drawback is often described in ITC technique and implied the need of very high quantities of enzymes and ligands to accurately determine a binding constant.<sup>310</sup> Further robust direct means techniques could be considered in the future to provide information on affinity, like the sensitive mass spectrometry or the label fluorescence methods. They are, however, time-consuming analyses with need of cautious conditions and may cause disruption in the synthesised cofactor (e.g. labelled experiments).

Indirect means were then applied to investigate the binding affinity. Using inhibition assays, smaller concentrations of synthesised catalysts were used to compete with the natural cofactor NAD(P)<sup>+</sup> inside ADHs. This method resulted in an IC<sub>50</sub> around low mM level for **49** and **50**, in agreement with the docking study. A higher affinity was obtained by the natural cofactor compared to the other ligands. However, low IC<sub>50</sub> values for the ligands still represented inhibition of the enzyme catalysis and gave insights into their affinity indirectly. In this project, the affinity for the enzyme is a main parameter for the design of AMs based on supramolecular anchoring. Therefore, a precise measurement of the binding affinity of the catalysts in ADHs must still be established.

The three iridium catalysts combined with the three ADHs were then evaluated for the imine reduction of 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline into the chiral amines 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline, in aqueous media. The three metal catalysts (**47**, **49** and **50**) tested alone exhibited catalytic activity without stereoselectivity, and the best turnover frequency was displayed by the catalyst **50**.

The data obtained for the catalyst-ADH constructs were more complicated. Catalyst **50** in combination with TbADH-5M showed reduction of cyclic imine without stereoselectivity. The two other catalysts and the wild-type scaffolds showed no activity. These results confirmed the inhibition assays' conclusion of a binding of all the catalysts inside the enzymes, with mutual inhibitions.

The theory that catalyst **50** bound correctly in the NADPH pocket was supported by the catalytic activity observed for the construct TbADH-5M - **50**. Besides, the lack of activity of the wild type enzymes in combination with **50** reinforced this theory. Indeed, the difference in the scaffolds was the presence of  $\text{Zn}^{2+}$  in the wild type, reducing the active site space and preventing accommodation of both catalyst **50** and substrate.

Experiments with His-HLADH lacking  $\text{Zn}^{2+}$  would be a plus to support this hypothesis. Besides, Michaelis Menten experiment on TbADH would confirm any binding inside a pocket, while the addition of NAD(P)H to the catalytic mixture could also conclude on a binding in the cofactor pocket if an inhibition of the activity of the TbADH-5M - **50** system would occur.

The lack of stereoselectivity was a major drawback and point of debate for the constructed AM. Absence of stereoselectivity has been explained in previous work with, for example, the exposition of the metal to the solvent and far from the chiral environment of the enzyme.<sup>79</sup> This would confirm some prior docking results which showed an orientation of the metal toward the pocket outside. A binding on the enzyme surface or a too large pocket, which would accommodate the ligand and the substrate, could also be considered.<sup>311</sup> However, a binding of catalyst **50** in these achiral positions would have also created active AMs with the wild type, which scaffold is similar to TbADH-5M.

Another explanation of the lack of stereoselectivity could be the need of refinement of the active site. At the beginning of their AM creation, Ward and co-workers also encountered very low stereoselectivity.<sup>67,57</sup> Further chemo-genetic optimisations of the ligands and the scaffold provided high selective AM.<sup>312</sup>

On the other hand, the presence of free catalyst could also explain activity with no stereoselectivity. The known binding of iridium metal to cysteine and histidine amino

acids<sup>301</sup> could explain the absence of activity for catalyst **50** with wild type enzymes, while the mutated TbADH-5M with no cysteine showed activity due to free catalyst **50** in solution. This hypothesis was even more suspected in His-HLADH by the presence of a chelating His-tag tail, known to bind cation transition metals. Ir(III) luminescent probes are notably used to selectivity bind histidine residues in proteins.<sup>313,314</sup> The metal binding theory is however debated by literature, where similar metals and enzymes have been used with success.<sup>301,307</sup> Besides, the hypothesis of activity from free catalyst **50** in solution was very unlikely, as the negative results of the other catalysts **47** and **49** with TbADH-5M and their low inhibition in competition with NADPH were again supporting the formation of an AM with catalyst **50**.

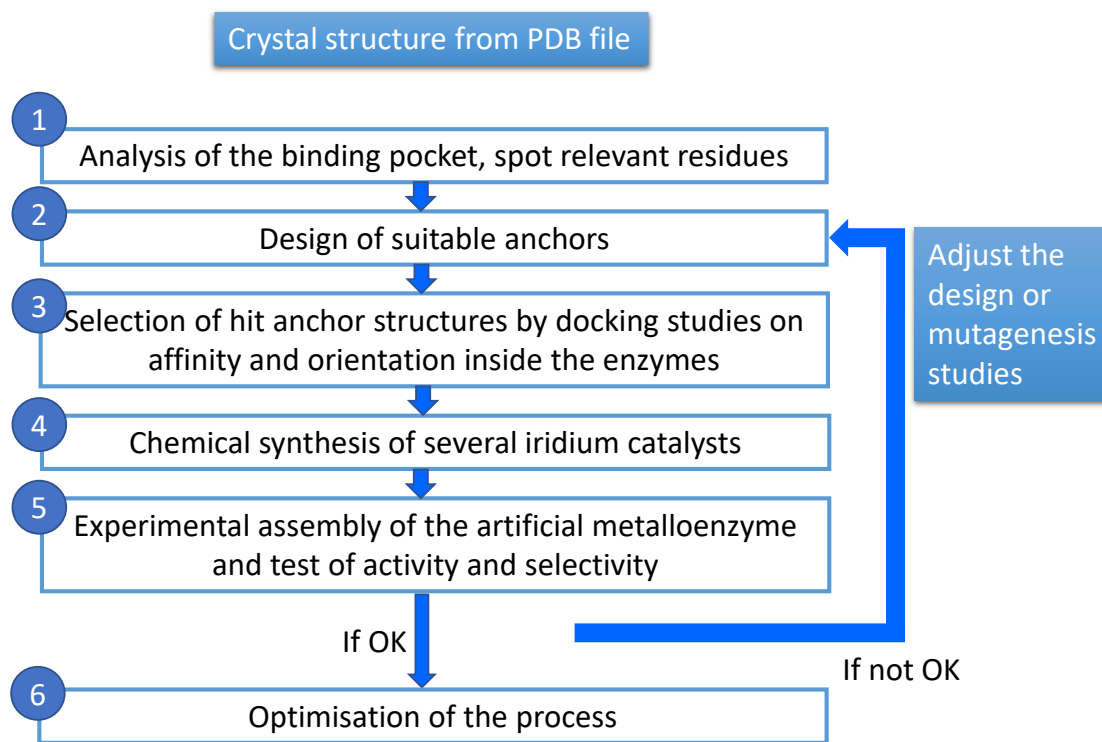
The absence of activity of catalysts **47** and **49** gave clear indirect evidence for the formation of an active AM with catalyst **50**. Their inhibition could then be explained by a different and unspecific binding of the anchoring part to the enzymes, in a non-reactive way. The main example would be a binding inside the substrate pocket, for example by the hydrophobic compound **49**. This theory was supported by the docking studies which predicted the insertion of ligand hydrophobic parts inside the substrate pocket. Future inhibition assays against the substrate would assess a binding inside the substrate pocket.

To investigate the creation of TbADH-5M - **50** AM and validate the docking approximations, further long-term approaches could be used. For example, NMR studies or crystallisation of this AM would provide rigorous details on the binding location of the metal complex and the spatial arrangement with the substrate.<sup>315</sup> The crystallisation of the other AMs developed might shed a light on the two other unspecific binding described.

Following the biochemical results, the better understanding of the anchor structure and the enzyme pocket will lead to future opportunities for substantial chemical and genetic optimisations (**Error! Reference source not found.**). The created supramolecular AM represented the beginning of an interesting concept, with only few enzymatic scaffolds previously used for the supramolecular engineering of AMs.<sup>45,47</sup>



By returning to the docking studies, screening for enzyme mutations and new anchors will result in discovering an increased affinity between the scaffold and the metal catalysts, ultimately resulting in a more reactive, robust, and selective AM.



**Figure 88:** Schematic representation of the rational design of an AM and the possible optimisations cycle.

The work presented in this thesis sets the foundations for future engineering of supramolecular AMs, capable of competing with current biological and chemical catalysts. The results proved the efficiency of a new rational design strategy using computational modelling in the conception of an active AM. As a renewable catalyst, this newly created AM participates in the development of green chemistry for the selective catalysis of green chemical reactions with high performance, mimicking natural products synthesis.

## 7 References

- 1 N. S. Scrutton, Speeding up enzyme engineering computationally, *IUCrJ.*, 2017, **4**, 5–6.
- 2 B. Hauer, Embracing Nature's Catalysts: A Viewpoint on the Future of Biocatalysis, *ACS Catal.*, 2020, **10**, 8418–8427.
- 3 H. S. Toogood and N. S. Scrutton, Enzyme engineering toolbox-A 'catalyst' for change, *Catal. Sci. Technol.*, 2013, **3**, 2182–2194.
- 4 J. C. Lewis, Artificial metalloenzymes and metallopeptide catalysts for organic synthesis, *ACS Catal.*, 2013, **3**, 2954–2975.
- 5 M. Dürrenberger and T. R. Ward, Recent achievements in the design and engineering of artificial metalloenzymes, *Curr. Opin. Chem. Biol.*, 2014, **19**, 99–106.
- 6 P. Kolb, R. S. Ferreira, J. J. Irwin and B. K. Shoichet, Docking and Chemoinformatic Screens for New Ligands and Targets, *Curr Opin Biotechnol.*, 2010, **20**, 429–436.
- 7 K. M. Elokely and R. J. Doerksen, Docking challenge: Protein sampling and molecular docking performance, *J. Chem. Inf. Model.*, 2013, **53**, 1934–1945.
- 8 J. Bos and G. Roelfes, Artificial metalloenzymes for enantioselective catalysis, *Curr. Opin. Chem. Biol.*, 2014, **19**, 135–143.
- 9 H. J. Davis and T. R. Ward, Artificial Metalloenzymes: Challenges and Opportunities, *ACS Cent. Sci.*, 2019, **5**, 1120–1136.
- 10 F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis and T. R. Ward, Artificial Metalloenzymes: Reaction Scope and Optimization Strategies, *Chem. Rev.*, 2018, **118**, 142–231.
- 11 F. Rosati and G. Roelfes, Artificial Metalloenzymes, *ChemCatChem*, 2010, **2**, 916–927.
- 12 M. Diéguez, J.-E. Bäckvall and O. Pàmies, *Artificial Metalloenzymes and MetalloDNAzymes in Catalysis*, 2018.
- 13 O. Pàmies, M. Diéguez and J.-E. Bäckvall, Artificial Metalloenzymes in Asymmetric Catalysis: Key Developments and Future Directions, *Adv. Synth. Catal.*, 2015, **357**, 1567–1586.
- 14 K. Yamamura and E. T. Kaiser, Studies on the oxidase activity of copper(II) carboxypeptidase A, *J. Chem. Soc. Chem. Commun.*, 1976, **20**, 830–831.
- 15 K. Okrasa and R. J. Kazlauskas, Manganese-substituted carbonic anhydrase as a new peroxidase, *Chem. A Eur. J.*, 2006, **12**, 1587–1596.
- 16 A. Fernández-Gacio, A. Codina, J. Fastrez, O. Riant and P. Soumillion, Transforming carbonic anhydrase into epoxide synthase by metal exchange,

*ChemBioChem*, 2006, **7**, 1013–1016.

- 17 T. Ueno, M. Suzuki, T. Goto, T. Matsumoto, K. Nagayama and Y. Watanabe, Size-selective olefin hydrogenation by a Pd nanocluster provided in an apo-ferritin cage, *Angew. Chemie Int. Ed.*, 2004, **43**, 2527–2530.
- 18 S. Abe, J. Niemeyer, M. Abe, Y. Takezawa, T. Ueno, T. Hikage, G. Erker and Y. Watanabe, Control of the coordination structure of organometallic palladium complexes in an apo-ferritin cage, *J. Am. Chem. Soc.*, 2008, **130**, 10512–10514.
- 19 N. Fujieda, A. Hasegawa, K. I. Ishihama and S. Itoh, Artificial dicopper oxidase: Rational reprogramming of bacterial metallo- $\beta$ -lactamase into a catechol oxidase, *Chem. Asian J.*, 2012, **7**, 1203–1207.
- 20 V. Köhler, J. Mao, T. Heinisch, A. Pordea, A. Sardo, Y. M. Wilson, L. Knörr, M. Creus, J. C. Prost, T. Schirmer and T. R. Ward, OsO<sub>4</sub>•streptavidin: A tunable hybrid catalyst for the enantioselective cis-dihydroxylation of olefins, *Angew. Chemie Int. Ed.*, 2011, **50**, 10863–10866.
- 21 J. Podtetenieff, A. Taglieber, E. Bill, E. J. Reijerse and M. T. Reetz, An artificial metalloenzyme: Creation of a designed copper binding site in a thermostable protein, *Angew. Chemie Int. Ed.*, 2010, **49**, 5151–5155.
- 22 I. Drienovská, A. Rioz-Martínez, A. Draksharapu and G. Roelfes, Novel artificial metalloenzymes by in vivo incorporation of metal-binding unnatural amino acids, *Chem. Sci.*, 2015, **6**, 770–776.
- 23 G. Roelfes, LmrR: A Privileged Scaffold for Artificial Metalloenzymes, *Acc. Chem. Res.*, 2019, **52**, 545–556.
- 24 A. Zanghellini, L. Jiang, A. M. Wollacott, G. Cheng, J. Meiler, E. A. Althoff, D. Röthlisberger and D. Baker, New algorithms and an in silico benchmark for computational enzyme design, *Protein Sci.*, 2006, **15**, 2785–2794.
- 25 J. H. Mills, S. D. Khare, J. M. Bolduc, F. Forouhar, V. K. Mulligan, S. Lew, J. Seetharaman, L. Tong, B. L. Stoddard and D. Baker, Computational design of an unnatural amino acid dependent metalloprotein with atomic level accuracy, *J. Am. Chem. Soc.*, 2013, **135**, 13393–13399.
- 26 H. L. Levine, Y. Nakagawa and E. T. Kaiser, Flavopapain: Synthesis and properties of semi-synthetic enzymes, *Biochem. Biophys. Res. Commun.*, 1977, **76**, 64–70.
- 27 R. R. Davies and M. D. Distefano, A semisynthetic metalloenzyme based on a protein cavity that catalyzes the enantioselective hydrolysis of ester and amide substrates, *J. Am. Chem. Soc.*, 1997, **119**, 11643–11652.
- 28 R. R. Davies, H. Kuang, D. Qi, A. Mazhary, E. Mayaan and M. D. Distefano, Artificial metalloenzymes based on protein cavities: Exploring the effect of altering the metal ligand attachment position by site directed mutagenesis, *Bioorganic Med. Chem. Lett.*, 1999, **9**, 79–84.
- 29 M. T. Reetz, M. Rentzsch, A. Pletsch and M. Maywald, Towards the directed

evolution of hybrid catalysts, *Chimia (Aarau).*, 2002, **56**, 721–723.

- 30 M. T. Reetz, M. Rentzsch, A. Pletsch, A. Taglieber, F. Hollman, R. J. G. Mondière, N. Dickman, B. Höcker, S. Cerrone, M. C. Hagaer and R. Sterner, A robust protein host for anchoring chelating ligands and organocatalysts, *ChemBioChem*, 2008, **9**, 552–564.
- 31 B. Talbi, P. Haquette, A. Martel, F. De Montigny, C. Fosse, S. Cordier, T. Roisnel, G. Jaouen and M. Salmain, ( $\eta^6$ -Arene) ruthenium(ii) complexes and metallo-papain hybrid as Lewis acid catalysts of Diels-Alder reaction in water, *Dalt. Trans.*, 2010, **39**, 5605–5607.
- 32 C. Zhang, P. Srivastava, K. Ellis-Guardiola and J. C. Lewis, Manganese terpyridine artificial metalloenzymes for benzylic oxygenation and olefin epoxidation, *Tetrahedron*, 2014, **70**, 4245–4249.
- 33 T. Himiyama, D. F. Sauer, A. Onoda, T. P. Spaniol, J. Okuda and T. Hayashi, Construction of a hybrid biocatalyst containing a covalently-linked terpyridine metal complex within a cavity of aponitrobindin, *J. Inorg. Biochem.*, 2016, **158**, 55–61.
- 34 C. Buron, K. Sénéchal-David, R. Ricoux, J. P. Le Caër, V. Guérineau, P. Méjanelle, R. Guillot, C. Herrero, J. P. Mahy and F. Banse, An Artificial Enzyme Made by Covalent Grafting of an FeII Complex into  $\beta$ -Lactoglobulin: Molecular Chemistry, Oxidation Catalysis, and Reaction-Intermediate Monitoring in a Protein, *Chem. A Eur. J.*, 2015, **21**, 12188–12193.
- 35 P. J. Deuss, G. Popa, C. H. Botting, W. Laan and P. C. J. Kamer, Highly efficient and site-selective phosphane modification of proteins through hydrazone linkage: Development of artificial metalloenzymes, *Angew. Chemie Int. Ed.*, 2010, **49**, 5315–5317.
- 36 W. Laan, B. K. Muñoz, R. Den Heeten and P. C. J. Kamer, Artificial metalloenzymes through cysteine-selective conjugation of phosphines to photoactive yellow protein, *ChemBioChem*, 2010, **11**, 1236–1239.
- 37 M. Basauri-Molina, D. G. A. Verhoeven, A. J. van Schaik, H. Kleijn and R. J. M. Klein Gebbink, Ring-Closing and Cross-Metathesis with Artificial Metalloenzymes Created by Covalent Active Site-Directed Hybridization of a Lipase, *Chemistry*, 2015, **21**, 15676–15685.
- 38 M. Basauri-Molina, C. F. Riemersma, M. A. Würdemann, H. Kleijn and R. J. M. Klein Gebbink, Lipase active site covalent anchoring of Rh(NHC) catalysts: Towards chemoselective artificial metalloenzymes, *Chem. Commun.*, 2015, **51**, 6792–6795.
- 39 D. K. Garner, L. Liang, D. A. Barrios, J. L. Zhang and Y. Lu, The important role of covalent anchor positions in tuning catalytic properties of a rationally designed Mnsalen-containing metalloenzyme, *ACS Catal.*, 2011, **1**, 1083–1089.
- 40 J. R. Carey, S. K. Ma, T. D. Pfister, D. K. Garne, H. K. Kim, J. A. Abramite, Z. Wang, Z. Guo and Y. Lu, A site-selective dual anchoring strategy for artificial

metalloprotein design, *J. Am. Chem. Soc.*, 2004, **126**, 10812–10813.

- 41 P. Srivastava, H. Yang, K. Ellis-Guardiola and J. C. Lewis, Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation, *Nat. Commun.*, 2015, **6**, 1–8.
- 42 J. Bos, F. Fusetti, A. J. M. Driessen and G. Roelfes, Enantioselective artificial metalloenzymes by creation of a novel active site at the protein dimer interface, *Angew. Chemie Int. Ed.*, 2012, **51**, 7472–7475.
- 43 J. Bos, A. García-Herraz and G. Roelfes, An enantioselective artificial metallo-hydratase, *Chem. Sci.*, 2013, **4**, 3578–3582.
- 44 F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum and V. L. Pecoraro, Protein design: toward functional metalloenzymes., *Chem. Rev.*, 2014, **114**, 3495–578.
- 45 M. Ohashi, T. Koshiyama, T. Ueno, M. Yanase, H. Fujii, Y. Watanabe, R. L. Osborne and J. H. Dawson, Preparation of artificial metalloenzymes by insertion of chromium (III) Schiff base complexes into apomyoglobin mutants, *Angew. Chem. Int. Ed.*, 2003, **42**, 1005–1008.
- 46 H. M. Key, P. Dydio, D. S. Clark and J. F. Hartwig, Abiological catalysis by artificial haem proteins containing noble metals in place of iron, *Nature*, 2016, **534**, 534–537.
- 47 P. Dydio, H. M. Key, H. Hayashi, D. S. Clark and J. F. Hartwig, Chemoselective, enzymatic C-H bond amination catalyzed by a cytochrome P450 containing an Ir(Me)-PIX cofactor, *J. Am. Chem. Soc.*, 2017, **139**, 1750–1753.
- 48 J. Bos, W. R. Browne, A. J. M. Driessen and G. Roelfes, Supramolecular Assembly of Artificial Metalloenzymes Based on the Dimeric Protein LmrR as Promiscuous Scaffold, *J. Am. Chem. Soc.*, 2015, **137**, 9796–9799.
- 49 Y. Okamoto and T. R. Ward, *Supramolecular Enzyme Mimics*, Elsevier, Second Edi., 2017, vol. 4.
- 50 M. E. Wilson and G. M. Whitesides, Conversion of a Protein to a Homogeneous Asymmetric Hydrogenation Catalyst by Site-Specific Modification with a Diphosphinerhodium(I) Moiety, *J. Am. Chem. Soc.*, 1978, **100**, 306–307.
- 51 C. C. Lin, C. W. Lin and A. S. C. Chan, Catalytic hydrogenation of itaconic acid in a biotinylated Pyrphos- rhodium(I) system in a protein cavity, *Tetrahedron Asymmetry*, 1999, **10**, 1887–1893.
- 52 T. K. Hyster and T. R. Ward, Genetic Optimization of Metalloenzymes: Enhancing Enzymes for Non-Natural Reactions, *Angew. Chemie - Int. Ed.*, 2016, **55**, 7344–7357.
- 53 J. Collot, J. Gradinaru, N. Humbert, M. Skander, A. Zocchi and T. R. Ward, Artificial metalloenzymes for enantioselective catalysis based on biotin-avidin, *J. Am. Chem. Soc.*, 2003, **125**, 9030–9031.
- 54 M. Skander, N. Humbert, J. Collot, J. Gradinaru, G. Klein, A. Loosli, J. Sauser, A.

- Zocchi, F. Gilardoni and T. R. Ward, Artificial metalloenzymes: (Strept)avidin as host for enantioselective hydrogenation by achiral biotinylated rhodium-diphosphine complexes, *J. Am. Chem. Soc.*, 2004, **126**, 14411–14418.
- 55 U. E. Rusbandi, C. Lo, M. Skander, A. Ivanova, M. Creus, N. Humbert and T. R. Ward, Second generation artificial hydrogenases based on the biotin-avidin technology: Improving activity, stability and selectivity by introduction of enantiopure amino acid spacers, *Adv. Synth. Catal.*, 2007, **349**, 1923–1930.
  - 56 R. Noyori and S. Hashiguchi, Asymmetric Transfer Hydrogenation Catalyzed by Chiral Ruthenium Complexes, *Acc. Chem. Res.*, 1997, **30**, 97–102.
  - 57 C. Letondor, N. Humbert and T. R. Ward, Artificial metalloenzymes based on biotin-avidin technology for the enantioselective reduction of ketones by transfer hydrogenation, *PNAS*, 2005, **102**, 4683–4687.
  - 58 C. Letondor, A. Pordea, N. Humbert, A. Ivanova, S. Mazurek, M. Novic and T. R. Ward, Artificial transfer hydrogenases based on the biotin-(strept)avidin technology: Fine tuning the selectivity by saturation mutagenesis of the host protein, *J. Am. Chem. Soc.*, 2006, **128**, 8320–8328.
  - 59 M. Creus, A. Pordea, T. Rossel, A. Sardo, C. Letondor, A. Ivanova, I. LeTrong, R. E. Stenkamp and T. R. Ward, X-ray structure and designed evolution of an artificial transfer hydrogenase, *Angew. Chemie Int. Ed.*, 2008, **47**, 1400–1404.
  - 60 C. M. Thomas, C. Letondor, N. Humbert and T. R. Ward, Aqueous oxidation of alcohols catalyzed by artificial metalloenzymes based on the biotin-avidin technology, *J. Organomet. Chem.*, 2005, **690**, 4488–4491.
  - 61 F. Schwizer, V. Köhler, M. Dürrenberger, L. Knörr and T. R. Ward, Genetic optimization of the catalytic efficiency of artificial imine reductases based on biotin-streptavidin technology, *ACS Catal.*, 2013, **3**, 1752–1755.
  - 62 M. Dürrenberger, T. Heinisch, Y. M. Wilson, T. Rossel, E. Nogueira, L. Knörr, A. Mutschler, K. Kersten, M. J. Zimbron, J. Pierron, T. Schirmer and T. R. Ward, Artificial transfer hydrogenases for the enantioselective reduction of cyclic imines, *Angew. Chemie Int. Ed.*, 2011, **50**, 3026–3029.
  - 63 T. Quinto, F. Schwizer, J. M. Zimbron, A. Morina, V. Köhler and T. R. Ward, Expanding the chemical diversity in artificial imine reductases based on the biotin-streptavidin technology, *ChemCatChem*, 2014, **6**, 1010–1014.
  - 64 M. Hesticová, T. Heinisch, L. Alonso-Cotchico, J. D. Maréchal, P. Vidossich and T. R. Ward, Directed Evolution of an Artificial Imine Reductase, *Angew. Chemie Int. Ed.*, 2018, **57**, 1863–1868.
  - 65 J. M. Zimbron, T. Heinisch, M. Schmid, D. Hamels, E. S. Nogueira, T. Schirmer and T. R. Ward, A dual anchoring strategy for the localization and activation of artificial metalloenzymes based on the biotin-streptavidin technology, *J. Am. Chem. Soc.*, 2013, **135**, 5384–5388.
  - 66 C. Lo, M. R. Ringenberg, D. Gndt, Y. Wilson and T. R. Ward, Artificial metalloenzymes for olefin metathesis based on the biotin-(strept)avidin

- technology, *Chem. Commun.*, 2011, **47**, 12065–12067.
- 67 A. Pordea, D. Mathis and T. R. Ward, Incorporation of biotinylated manganese-salen complexes into streptavidin: New artificial metalloenzymes for enantioselective sulfoxidation, *J. Organomet. Chem.*, 2009, **694**, 930–936.
  - 68 A. Chatterjee, H. Mallin, J. Klehr, J. Vallapurackal, A. D. Finke, L. Vera, M. Marsh and T. R. Ward, An enantioselective artificial Suzukiase based on the biotin-streptavidin technology, *Chem. Sci.*, 2016, **7**, 673–677.
  - 69 T. K. Hyster, L. Knörr, T. R. Ward and T. Rovis, Biotinylated Rh(III) Complexes in Engineered Streptavidin for Accelerated Asymmetric C–H Activation, *Science*, 2012, **338**, 500–503.
  - 70 V. Köhler, Y. M. Wilson, M. Dürrenberger, D. Ghislieri, E. Churakova, T. Quinto, L. Knörr, D. Häussinger, F. Hollmann, N. J. Turner and T. R. Ward, Synthetic cascades are enabled by combining biocatalysts with artificial metalloenzymes, *Nat. Chem.*, 2013, **5**, 93–99.
  - 71 M. D. Truppo, in *Comprehensive Chirality*, Elsevier Ltd., 2012, vol. 7, pp. 46–70.
  - 72 H. Wu, C. Tian, X. Song, C. Liu, D. Yang and Z. Jiang, Methods for the regeneration of nicotinamide coenzymes, *Green Chem.*, 2013, **15**, 1773–1789.
  - 73 Y. Okamoto, V. Köhler, C. E. Paul, F. Hollmann and T. R. Ward, Efficient in situ regeneration of NADH mimics by an artificial metalloenzyme, *ACS Catal.*, 2016, **6**, 3553–3557.
  - 74 Y. Okamoto, V. Köhler and T. R. Ward, An NAD(P)H-Dependent artificial transfer hydrogenase for multienzymatic cascades, *J. Am. Chem. Soc.*, 2016, **138**, 5781–5784.
  - 75 A. D. Liang, J. Serrano-Plana, R. L. Peterson and T. R. Ward, Artificial Metalloenzymes Based on the Biotin-Streptavidin Technology: Enzymatic Cascades and Directed Evolution, *Acc. Chem. Res.*, 2019, **52**, 585–595.
  - 76 Y. Okamoto, R. Kojima, F. Schwizer, E. Bartolami, T. Heinisch, S. Matile, M. Fussenegger and T. R. Ward, A cell-penetrating artificial metalloenzyme regulates a gene switch in a designer mammalian cell, *Nat. Commun.*, 2018, **9**, 1–7.
  - 77 M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, Directed evolution of artificial metalloenzymes for in vivo metathesis, *Nature*, 2016, **537**, 661–665.
  - 78 T. Heinisch, F. Schwizer, B. Garabedian, E. Csibra, M. Jeschek, J. Vallapurackal, V. B. Pinheiro, P. Marlière, S. Panke and T. R. Ward, E. coli surface display of streptavidin for directed evolution of an allylic deallylase, *Chem. Sci.*, 2018, **9**, 5383–5388.
  - 79 E. Sansiaume-Dagousset, A. Urvoas, K. Chelly, W. Ghattas, J. D. Maréchal, J. P. Mahy and R. Ricoux, Neocarzinostatin-based hybrid biocatalysts for oxidation reactions, *Dalt. Trans.*, 2014, **43**, 8344–8354.

- 80 A. Urvoas, W. Ghattas, J. D. Maréchal, F. Avenier, F. Bellande, W. Mao, R. Ricoux and J. P. Mahy, Neocarzinostatin-based hybrid biocatalysts with a RNase like activity, *Bioorganic Med. Chem.*, 2014, **22**, 5678–5686.
- 81 W. Ghattas, L. Cotchico-Alonso, J. D. Maréchal, A. Urvoas, M. Rousseau, J. P. Mahy and R. Ricoux, Artificial Metalloenzymes with the Neocarzinostatin Scaffold: Toward a Biocatalyst for the Diels-Alder Reaction, *ChemBioChem*, 2016, **17**, 433–440.
- 82 F. W. Monnard, T. Heinisch, E. S. Nogueira, T. Schirmer and T. R. Ward, Human Carbonic Anhydrase II as a host for piano-stool complexes bearing a sulfonamide anchor, *Chem. Commun.*, 2011, **47**, 8238–8240.
- 83 F. W. Monnard, E. S. Nogueira, T. Heinisch, T. Schirmer and T. R. Ward, Human carbonic anhydrase II as host protein for the creation of artificial metalloenzymes: The asymmetric transfer hydrogenation of imines, *Chem. Sci.*, 2013, **4**, 3269–3274.
- 84 T. Heinisch, M. Pellizzoni, M. Dürrenberger, C. E. Tinberg, V. Köhler, J. Klehr, D. Häussinger, D. Baker and T. R. Ward, Improving the catalytic performance of an artificial metalloenzyme by computational design, *J. Am. Chem. Soc.*, 2015, **137**, 10414–10419.
- 85 J. Zhao, A. Kajetanowicz and T. R. Ward, Carbonic anhydrase II as host protein for the creation of a biocompatible artificial metathesase, *Org. Biomol. Chem.*, 2015, **13**, 5652–5655.
- 86 S. Eda, I. Nasibullin, K. Vong, N. Kudo, M. Yoshida, A. Kurbangalieva and K. Tanaka, Biocompatibility and therapeutic potential of glycosylated albumin artificial metalloenzymes, *Nat. Catal.*, 2019, **2**, 780–792.
- 87 F. Hollmann, I. W. C. E. Arends and D. Holtmann, Enzymatic reductions for the chemist, *Green Chem.*, 2011, **13**, 2285–2313.
- 88 K. L. Kavanagh, H. Jörnvall, B. Persson and U. Oppermann, Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes, *Cell. Mol. Life Sci.*, 2008, **65**, 3895–3906.
- 89 C.-T. Eklund, H. Nordtrom, B. Zeppezauer, E. Soderlund, G. Ohlsson, I. Boiwe, T. Soderberg B-O., Tapia, O. Branden, Three-dimensional Structure of Horse Liver Alcohol Dehydrogenase at 2.4 Å Resolution, *J. Mol. Biol.*, 1976, **102**, 27–59.
- 90 Y. Yang and H. Zhou, Effect of zinc ions on conformational stability of yeast alcohol dehydrogenase, *Biokhimiya*, 2001, **66**, 61–70.
- 91 A. Dołęga, Alcohol dehydrogenase and its simple inorganic models, *Coord. Chem. Rev.*, 2010, **254**, 916–937.
- 92 H. Eklund and S. Ramaswamy, Medium- and short-chain dehydrogenase/reductase gene and protein families: Three-dimensional structures of MDR alcohol dehydrogenases, *Cell. Mol. Life Sci.*, 2008, **65**, 3907–



3917.

- 93 C. Lee, D. L. Bedgar, L. B. Davin and N. G. Lewis, Assessment of a putative proton relay in Arabidopsis cinnamyl alcohol dehydrogenase catalysis, *Org. Biomol. Chem.*, 2013, **11**, 1127–1134.
- 94 F. Hollmann, I. W. C. E. Arends and D. Holtmann, Enzymatic reductions for the chemist, *Green Chem.*, 2011, **13**, 2285–2313.
- 95 D. Ji, L. Wang, S. Hou, W. Liu, J. Wang, Q. Wang and Z. K. Zhao, Creation of bioorthogonal redox systems depending on nicotinamide flucytosine dinucleotide, *J. Am. Chem. Soc.*, 2011, **133**, 20857–20862.
- 96 B. Persson, J. Hedlund and H. Jörnvall, Medium- and short-chain dehydrogenase/reductase gene and protein families: The MDR superfamily, *Cell. Mol. Life Sci.*, 2008, **65**, 3879–3894.
- 97 Y. Korkhin, A. J. Kalb, M. Peretz, O. Bogin, Y. Burstein and F. Frolov, NADP-dependent bacterial alcohol dehydrogenases: Crystal structure, cofactor-binding and cofactor specificity of the ADHs of *Clostridium beijerinckii* and *Thermoanaerobacter brockii*, *J. Mol. Biol.*, 1998, **278**, 967–981.
- 98 F. Fan, J. A. Lorenzen and B. V. Plapp, An aspartate residue in yeast alcohol dehydrogenase I determines the specificity for coenzyme., *Biochemistry*, 1991, **30**, 6397–6401.
- 99 I. Fita, A. Rosell, E. Valencia, X. Pare and W. F. Ochoa, Crystal Structure of the Vertebrate NADP ( H ) - dependent Alcohol Dehydrogenase ( ADH8 ), *J. Mol. Biol.*, 2003, **330**, 75–85.
- 100 R. Lamed and J. G. Zeikus, Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in *Clostridium thermocellum* and *Thermoanaerobium brockii*, *J. Bacteriol.*, 1980, **144**, 569–578.
- 101 M. Peretz and Y. Burstein, Amino Acid Sequence of Alcohol Dehydrogenase from the Thermophilic Bacterium *Thermoanaerobium brockii*, *Biochemistry*, 1989, **28**, 6549–6555.
- 102 T. Quinto, V. Köhler and T. R. Ward, Recent trends in biomimetic NADH regeneration, *Top. Catal.*, 2014, **57**, 321–331.
- 103 R. J. Ansell, D. A. P. Small and C. R. Lowe, The interactions of artificial coenzymes with alcohol dehydrogenase and other NAD(P)(H)/dependent enzymes, *J. Mol. Catal.*, 1999, **6**, 111–123.
- 104 H. C. Lo and R. H. Fish, Biomimetic NAD<sup>+</sup> Models for Tandem Cofactor Regeneration, Horse Liver Alcohol Dehydrogenase Recognition of 1,4-NADH Derivatives, and Chiral Synthesis, *Angew. Chemie Int. Ed.*, 2002, **41**, 478–481.
- 105 H. C. Lo, J. D. Ryan, J. B. Kerr, D. S. Clark and R. H. Fish, Bioorganometallic chemistry: Co-factor regeneration, enzyme recognition of biomimetic 1,4-NADH analogs, and organic synthesis; tandem catalyzed regioselective

- formation of N-substituted-1,4- dihydronicotinamide derivatives with  $[\text{Cp}^*\text{Rh}(\text{bpy})\text{H}]^+$ , coupled to chiral S-alcohol formation with HLADH, and engineered cytochrome P450s, for selective C-H oxidation reactions, *J. Organomet. Chem.*, 2017, **839**, 1–15.
- 106 C. E. Paul and F. Hollmann, A survey of synthetic nicotinamide cofactors in enzymatic processes, *Appl. Microbiol. Biotechnol.*, 2016, **100**, 4773–4778.
  - 107 T. Knaus, C. E. Paul, C. W. Levy, S. De Vries, F. G. Mutti, F. Hollmann and N. S. Scrutton, Better than Nature: Nicotinamide Biomimetics That Outperform Natural Coenzymes, *J. Am. Chem. Soc.*, 2016, **138**, 1033–1039.
  - 108 C. Nowak, A. Pick, P. Lommes and V. Sieber, Enzymatic Reduction of Nicotinamide Biomimetic Cofactors Using an Engineered Glucose Dehydrogenase: Providing a Regeneration System for Artificial Cofactors, *ACS Catal.*, 2017, **7**, 5202–5208.
  - 109 M. Basle, Development of an artificial metalloenzyme from a nicotinamide-dependent enzymatic scaffold, *Univ. Nottingham*, 2019, 1–217.
  - 110 J. G. De Vries and N. Mršić, Organocatalytic asymmetric transfer hydrogenation of imines, *Catal. Sci. Technol.*, 2011, **1**, 727–735.
  - 111 B. Knoevenagel, E. Bergdolt, Ueber das Verhalten des L/2.5-Dihydroterephthalsäuredimethylesters bei höheren Temperaturen und in Gegenwart von Palladiummohr., *Berichte der Dtsch. Chem. Gesellschaft*, 1903, **36**, 2857–2860.
  - 112 R. Meerwein, H.; Schmidt, Ein neues Verfahren zur Reduktion von Aldehyden und Ketonen., *Justus Liebigs Ann. Chem.*, 1925, **444**, 221–238.
  - 113 A. Verley, Sur l'échange de groupements fonctionnels entre deux molecules. Passage de la fonction alcool a la fonction aldehyde et inversement., *Bull. Soc. Chim. Fr*, 1925, **37**, 537–542.
  - 114 T. R. B. Haddad, Y.M.Y.; Henbest, H.B.; Husbands, J.; Mitchell, Reduction of Cyclohexanones to Axial Alcohols via Iridium-containing Catalysts, *Proc. Chem. Soc.*, 1964, 361–361.
  - 115 Y. Sasson and J. Blum, Homogeneous catalytic transfer-hydrogenation of  $\alpha,\beta$ -unsaturated carbonyl compounds by dichlorotris(triphenylphosphine)ruthenium (II), *Tetrahedron Lett.*, 1971, **12**, 2167–2170.
  - 116 Y. Sasson and J. Blum, Dichlorotris(triphenylphosphine)ruthenium-Catalyzed Hydrogen Transfer from Alcohols to Saturated and  $\alpha,\beta$ -Unsaturated Ketones, *J. Org. Chem.*, 1975, **40**, 1887–1896.
  - 117 W. S. Knowles, Asymmetric Hydrogenations (Nobel Lecture 2001), *Angew. Chemie Int. Ed.*, 2002, **41**, 1998–2007.
  - 118 R. Noyori, Asymmetric catalysis: Science and opportunities (nobel lecture), *Angew. Chemie Int. Ed.*, 2002, **41**, 2008–2022.

- 119 C. Wang, B. Villa-Marcos and J. Xiao, Hydrogenation of imino bonds with half-sandwich metal catalysts, *Chem. Commun.*, 2011, **47**, 9773–9785.
- 120 E. Baráth, Hydrogen transfer reactions of carbonyls, Alkynes, and Alkenes with noble metals in the presence of alcohols/ethers and amines as hydrogen donors, *Catalysts*, 2018, **8**, 1–25.
- 121 N. Fleury-Brégeot, V. de la Fuente, S. Castillón and C. Claver, Highlights of transition metal-catalyzed asymmetric hydrogenation of imines, *ChemCatChem*, 2010, **2**, 1346–1371.
- 122 L. Wang, Q. Zhou, C. Qu, Q. Wang, L. Cun, J. Zhu and J. Deng, Efficient asymmetric transfer hydrogenation of N-sulfonylimines on water, *Tetrahedron*, 2013, **69**, 6500–6506.
- 123 N. Haraguchi, K. Tsuru, Y. Arakawa and S. Itsuno, Asymmetric transfer hydrogenation of imines catalyzed by a polymer-immobilized chiral catalyst, *Org. Biomol. Chem.*, 2009, **7**, 69–75.
- 124 J. Canivet and G. Süß-Fink, Water-soluble arene ruthenium catalysts containing sulfonated diamine ligands for asymmetric transfer hydrogenation of  $\alpha$ -aryl ketones and imines in aqueous solution, *Green Chem.*, 2007, **9**, 391–397.
- 125 C. Wang, C. Li, X. Wu, A. Pettman and J. Xiao, PH-regulated asymmetric transfer hydrogenation of quinolines in water, *Angew. Chemie Int. Ed.*, 2009, **48**, 6524–6528.
- 126 M. Soetens, F. Drouet and O. Riant, ( $\eta^5$ -Pentamethylcyclopentadienyl)iridium Complex Catalyzed Imine Reductions Utilizing the Biomimetic 1,4-NAD(P)H Cofactor and N-Benzyl-1,4-dihydronicotinamide as the Hydride-Transfer Agent, *ChemCatChem*, 2017, **9**, 929–933.
- 127 J. P. C. Coverdale, I. Romero-Canelón, C. Sanchez-Cano, G. J. Clarkson, A. Habtemariam, M. Wills and P. J. Sadler, Asymmetric transfer hydrogenation by synthetic catalysts in cancer cells, *Nat. Chem.*, 2018, **10**, 347–354.
- 128 J. W. Yang, M. T. Hechavarria Fonseca and B. List, A metal-free transfer hydrogenation: Organocatalytic conjugate reduction of  $\alpha,\beta$ -unsaturated aldehydes, *Angew. Chemie Int. Ed.*, 2004, **43**, 6660–6662.
- 129 S. G. Ouellet, J. B. Tuttle and D. W. C. MacMillan, Enantioselective organocatalytic hydride reduction, *J. Am. Chem. Soc.*, 2005, **127**, 32–33.
- 130 S. Mayer and B. List, Asymmetric counteranion-directed catalysis, *Angew. Chemie Int. Ed.*, 2006, **45**, 4193–4195.
- 131 A. Michrowska and B. List, Concise synthesis of ricciocarpin A and discovery of a more potent analogue, *Nat. Chem.*, 2009, **1**, 225–228.
- 132 S. G. Ouellet, A. M. Walji and D. W. C. Macmillan, Enantioselective organocatalytic transfer hydrogenation reactions using Hantzsch esters, *Acc. Chem. Res.*, 2007, **40**, 1327–1339.

- 133 J. W. Yang and B. List, Catalytic asymmetric transfer hydrogenation of  $\alpha$ -ketoesters with hantzsch esters, *Org. Lett.*, 2006, **8**, 5653–5655.
- 134 L. Zhang, J. Yuan, Y. Xu, Y. P. Zhang and X. Qian, New artificial fluoro-cofactor of hydride transfer with novel fluorescence assay for redox, *Chem. Commun.*, 2016, **52**, 6471–6474.
- 135 A. Chevalley, M. V. Cherrier, J. C. Fontecilla-Camps, M. Ghasemi and M. Salmain, Artificial metalloenzymes derived from bovine  $\beta$ -lactoglobulin for the asymmetric transfer hydrogenation of an aryl ketone-synthesis, characterization and catalytic activity, *Dalt. Trans.*, 2014, **43**, 5482–5489.
- 136 N. Madern, N. Queyriaux, A. Chevalley, M. Ghasemi, O. Nicolotti, I. Ciofini, G. F. Mangiatordi and M. Salmain, Piano-stool d6-rhodium(III) complexes of chelating pyridine-based ligands and their papain bioconjugates for the catalysis of transfer hydrogenation of aryl ketones in aqueous medium, *J. Mol. Catal. B Enzym.*, 2015, **122**, 314–322.
- 137 T. C. Nugent, *Chiral Amine Synthesis: Methods, Developments and Applications*, 2010.
- 138 K. Mitsukura, M. Suzuki, S. Shinoda, T. Kuramoto, T. Yoshida and T. Nagasawa, Purification and characterization of a novel (R)-imine reductase from *Streptomyces* sp. GF3587, *Biosci. Biotechnol. Biochem.*, 2011, **75**, 1778–1782.
- 139 P. N. Scheller, S. Fademrecht, S. Hofelzer, J. Pleiss, F. Leipold, N. J. Turner, B. M. Nestl and B. Hauer, Enzyme Toolbox: Novel Enantiocomplementary Imine Reductases, *ChemBioChem*, 2014, **15**, 2201–2204.
- 140 J. R. Marshall, P. Yao, S. L. Montgomery, J. D. Finnigan, T. W. Thorpe, R. B. Palmer, J. Mangas-Sanchez, R. A. M. Duncan, R. S. Heath, K. M. Graham, D. J. Cook, S. J. Charnock and N. J. Turner, Screening and characterization of a diverse panel of metagenomic imine reductases for biocatalytic reductive amination, *Nat. Chem.*, 2021, **13**, 140–148.
- 141 S. Hussain, F. Leipold, H. Man, E. Wells, S. P. France, K. R. Mulholland, G. Grogan and N. J. Turner, An (R)-imine reductase biocatalyst for the asymmetric reduction of cyclic imines, *ChemCatChem*, 2015, **7**, 579–583.
- 142 G. Grogan and N. J. Turner, InspiRED by Nature: NADPH-Dependent Imine Reductases (IREDS) as Catalysts for the Preparation of Chiral Amines, *Chem. - A Eur. J.*, 2016, **22**, 1900–1907.
- 143 G. Facchetti and I. Rimoldi, 8-Amino-5,6,7,8-tetrahydroquinoline in iridium(III) biotinylated Cp\* complex as artificial imine reductase, *New J. Chem.*, 2018, **42**, 18773–18776.
- 144 M. Pellizzoni, G. Facchetti, R. Gandolfi, M. Fusè, A. Contini and I. Rimoldi, Evaluation of Chemical Diversity of Biotinylated Chiral 1,3-Diamines as a Catalytic Moiety in Artificial Imine Reductase, *ChemCatChem*, 2016, **8**, 1665–1670.
- 145 A.-K. Raines, D. J. Clarke, J. E. Blagova, E. V. Dodson, E. J. Wilson, K. S. Duhme-

- Klair, Redox-switchable siderophore anchor enables reversible artificial metalloenzyme assembly., *Nat. Catal.*, 2018, **1**, 680–688.
- 146 B. A. Amrein, F. Steffen-Munsberg, I. Szeler, M. Purg, Y. Kulkarni and S. C. L. Kamerlin, CADEE: Computer-Aided Directed Evolution of Enzymes, *IUCrJ.*, 2017, **4**, 50–64.
  - 147 G. A. Strohmeier, H. Pichler, O. May and M. Gruber-Khadjawi, Application of designed enzymes in organic synthesis, *Chem. Rev.*, 2011, **111**, 4141–4164.
  - 148 I. D. Petrik, J. Liu and Y. Lu, Metalloenzyme design and engineering through strategic modifications of native protein scaffolds, *Curr. Opin. Chem. Biol.*, 2014, **19**, 67–75.
  - 149 S. Lutz, Beyond directed evolution-semi-rational protein engineering and design, *Curr. Opin. Biotechnol.*, 2010, **21**, 734–743.
  - 150 R. Li, H. J. Wijma, L. Song, Y. Cui, M. Otzen, Y. Tian, J. Du, T. Li, D. Niu, Y. Chen, J. Feng, J. Han, H. Chen, Y. Tao, D. B. Janssen and B. Wu, Computational redesign of enzymes for regio- and enantioselective hydroamination article, *Nat. Chem. Biol.*, 2018, **14**, 664–670.
  - 151 K. D. Miner, A. Mukherjee, Y. G. Gao, E. L. Null, I. D. Petrik, X. Zhao, N. Yeung, H. Robinson and Y. Lu, A designed functional metalloenzyme that reduces O<sub>2</sub> to H<sub>2</sub>O with over one thousand turnovers, *Angew. Chemie Int. Ed.*, 2012, **51**, 5589–5592.
  - 152 L. Le Li, H. Yuan, F. Liao, B. He, S. Q. Gao, G. B. Wen, X. Tan and Y. W. Lin, Rational design of artificial dye-decolorizing peroxidases using myoglobin by engineering Tyr/Trp in the heme center, *Dalt. Trans.*, 2017, **46**, 11230–11238.
  - 153 S. D. Khare, Y. Kipnis, P. J. Greisen, R. Takeuchi, Y. Ashani, M. Goldsmith, Y. Song, J. L. Gallaher, I. Silman, H. Leader, J. L. Sussman, B. L. Stoddard, D. S. Tawfik and D. Baker, Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis, *Nat. Chem. Biol.*, 2012, **8**, 294–300.
  - 154 H. S. Park, S. H. Nam, J. K. Lee, C. N. Yoon, B. Mannervik, S. J. Benkovic and H. S. Kim, Design and evolution of new catalytic activity with an existing protein scaffold, *Science*, 2006, **311**, 535–538.
  - 155 J. K. Leman, B. D. Weitzner, S. M. Lewis, J. Adolf-Bryfogle, N. Alam, R. F. Alford, M. Aprahamian, D. Baker, K. A. Barlow, P. Barth, B. Basanta, B. J. Bender, K. Blacklock, J. Bonet, S. E. Boyken, P. Bradley, C. Bystroff, P. Conway, S. Cooper, B. E. Correia, B. Coventry, R. Das, R. M. De Jong, F. DiMaio, L. Dsilva, R. Dunbrack, A. S. Ford, B. Frenz, D. Y. Fu, C. Geniesse, L. Goldschmidt, R. Gowthaman, J. J. Gray, D. Gront, S. Guffy, S. Horowitz, P. S. Huang, T. Huber, T. M. Jacobs, J. R. Jeliakov, D. K. Johnson, K. Kappel, J. Karanicolas, H. Khakzad, K. R. Khar, S. D. Khare, F. Khatib, A. Khramushin, I. C. King, R. Kleffner, B. Koepnick, T. Kortemme, G. Kuenze, B. Kuhlman, D. Kuroda, J. W. Labonte, J. K. Lai, G. Lapidoth, A. Leaver-Fay, S. Lindert, T. Linsky, N. London, J. H. Lubin, S. Lyskov, J. Maguire, L. Malmström, E. Marcos, O. Marcu, N. A. Marze, J. Meiler,

- R. Moretti, V. K. Mulligan, S. Nerli, C. Norn, S. Ó'Conchúir, N. Ollikainen, S. Ovchinnikov, M. S. Pacella, X. Pan, H. Park, R. E. Pavlovicz, M. Pethe, B. G. Pierce, K. B. Pilla, B. Raveh, P. D. Renfrew, S. S. R. Burman, A. Rubenstein, M. F. Sauer, A. Scheck, W. Schief, O. Schueler-Furman, Y. Sedan, A. M. Sevy, N. G. Sgourakis, L. Shi, J. B. Siegel, D. A. Silva, S. Smith, Y. Song, A. Stein, M. Szegedy, F. D. Teets, S. B. Thyme, R. Y. R. Wang, A. Watkins, L. Zimmerman and R. Bonneau, Macromolecular modeling and design in Rosetta: recent methods and frameworks, *Nat. Methods*, 2020, **17**, 665–680.
- 156 S. Studer, D. A. Hansen, Z. L. Pianowski, P. R. E. Mittl, A. Debon, S. L. Guffy, B. S. Der, B. Kuhlman and D. Hilvert, Evolution of a highly active and enantiospecific metalloenzyme from short peptides, *Science*, 2018, **362**, 1285–1288.
  - 157 S. Basler, S. Studer, Y. Zou, T. Mori, Y. Ota, A. Camus, H. A. Bunzel, R. C. Helgeson, K. N. Houk, G. Jiménez-Osés and D. Hilvert, Efficient Lewis acid catalysis of an abiological reaction in a de novo protein scaffold, *Nat. Chem.*, 2021, **13**, 231–235.
  - 158 J. Damborsky and J. Brezovsky, Computational tools for designing and engineering enzymes, *Curr. Opin. Chem. Biol.*, 2014, **19**, 8–16.
  - 159 A. Roy, A. Kucukural and Y. Zhang, I-TASSER: A unified platform for automated protein structure and function prediction, *Nat. Protoc.*, 2010, **5**, 725–738.
  - 160 A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Židek, A. W. R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D. T. Jones, D. Silver, K. Kavukcuoglu and D. Hassabis, Improved protein structure prediction using potentials from deep learning, *Nature*, 2020, **577**, 706–710.
  - 161 P. Kolb, R. S. Ferreira, J. J. Irwin and B. K. Shoichet, Docking and chemoinformatic screens for new ligands and targets, *Curr. Opin. Biotechnol.*, 2009, **20**, 429–436.
  - 162 I. A. Guedes, C. S. de Magalhães and L. E. Dardenne, Receptor-ligand molecular docking, *Biophys. Rev.*, 2014, **6**, 75–87.
  - 163 S. F. Sousa, A. J. M. Ribeiro, J. T. S. Coimbra, R. P. P. Neves, S. A. Martins, N. S. H. N. Moorthy, P. A. Fernandes and M. J. Ramos, Protein-Ligand Docking in the New Millennium – A Retrospective of 10 Years in the Field, *Curr. Med. Chem.*, 2013, **20**, 2296–2314.
  - 164 I. A. Guedes, C. S. de Magalhaes and L. E. Dardenne, Receptor-ligand molecular docking, *Biophys. Rev.*, 2014, **6**, 75–87.
  - 165 Y. Deng and B. Roux, Computations of standard binding free energies with molecular dynamics simulations, *J. Phys. Chem. B*, 2009, **113**, 2234–2246.
  - 166 H. Oshima, S. Re and Y. Sugita, Prediction of Protein-Ligand Binding Pose and Affinity Using the gREST+FEP Method, *J. Chem. Inf. Model.*, 2020, **60**, 5382–5394.

- 167 V. Muñoz Robles, E. Ortega-Carrasco, L. Alonso-Cotchico, J. Rodríguez-Guerra, A. Lledós and J. D. Maréchal, Toward the computational design of artificial metalloenzymes: From protein-ligand docking to multiscale approaches, *ACS Catal.*, 2015, **5**, 2469–2480.
- 168 S. Lutz, Beyond directed evolution - semi-rational protein engineering and design, *Curr. Opin. Biotechnol.*, 2011, **21**, 734–743.
- 169 A. Maršavelski, D. Petrović, P. Bauer, R. Vianello and S. C. L. Kamerlin, Empirical valence bond simulations suggest a direct hydride transfer mechanism for human diamine oxidase, *ACS Omega*, 2018, **3**, 3665–3674.
- 170 C. M. Bathelt, A. J. Mulholland and J. N. Harvey, QM/MM modeling of benzene hydroxylation in human cytochrome P450 2C9, *J. Phys. Chem. A*, 2008, **112**, 13149–13156.
- 171 S. M. Marques, D. Bednar and J. Damborsky, Computational study of protein-ligand unbinding for enzyme engineering, *Front. Chem.*, 2019, **6**, 1–15.
- 172 R. Ricoux, R. Dubuc, C. Dupont, J. D. Marechal, A. Martin, M. Sellier and J. P. Mahy, Hemozymes peroxidase activity of artificial hemoproteins constructed from the *Streptomyces lividans* xylanase A and iron(III)-carboxy-substituted porphyrins, *Bioconjug. Chem.*, 2008, **19**, 899–910.
- 173 L. Alonso-Cotchico, J. Rodríguez-Guerra, A. Lledós and J.-D. Maréchal, Molecular Modeling for Artificial Metalloenzyme Design and Optimization, *Acc. Chem. Res.*, 2020, **53**, 896–905.
- 174 J. E. Sánchez-Aparicio, G. Sciortino, D. V. Herrmannsdoerfer, P. O. Chueca, J. R. G. Pedregal and J. D. Maréchal, Gpathfinder: Identification of ligand-binding pathways by a multi-objective genetic algorithm, *Int. J. Mol. Sci.*, 2019, **20**, 3155.
- 175 E. Ortega-Carrasco, A. Lledós and J. D. Maréchal, Unravelling novel synergies between organometallic and biological partners: A quantum mechanics/molecular mechanics study of an artificial metalloenzyme, *J. R. Soc. Interface*, 2014, **11**, 1–9.
- 176 V. Muñoz Robles, P. Vidossich, A. Lledós, T. R. Ward and J. D. Maréchal, Computational insights on an artificial imine reductase based on the biotin-streptavidin technology, *ACS Catal.*, 2014, **4**, 833–842.
- 177 M. Allard, C. Dupont, V. Muñoz Robles, N. Doucet, A. Lledós, J. D. Maréchal, A. Urvoas, J. P. Mahy and R. Ricoux, Incorporation of Manganese Complexes into Xylanase: New Artificial Metalloenzymes for Enantioselective Epoxidation, *ChemBioChem*, 2012, **13**, 240–251.
- 178 C. Esmieu, M. V. Cherrier, P. Amara, E. Girgenti, C. Marchi-Delapierre, F. Oddon, M. Iannello, A. Jorge-Robin, C. Cavazza and S. Ménage, An artificial oxygenase built from scratch: Substrate binding site identified using a docking approach, *Angew. Chemie Int. Ed.*, 2013, **52**, 3922–3925.
- 179 H. Kries, R. Blomberg and D. Hilvert, De novo enzymes by computational

- design, *Curr. Opin. Chem. Biol.*, 2013, **17**, 221–228.
- 180 X. Zhang, J. DeChancie, H. Gunaydin, A. B. Chowdry, F. R. Clemente, A. J. T. Smith, T. M. Handel and K. N. Houk, Quantum mechanical design of enzyme active sites, *J. Org. Chem.*, 2008, **73**, 889–899.
  - 181 J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. St.Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael and D. Baker, Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction, *Science*, 2010, **329**, 309–313.
  - 182 D. N. Bolon and S. L. Mayo, Enzyme-like proteins by computational design, *PNAS*, 2001, **98**, 14274–14279.
  - 183 D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik and D. Baker, Kemp elimination catalysts by computational enzyme design, *Nature*, 2008, **453**, 190–195.
  - 184 S. Bjelic, L. G. Nivón, N. Çelebi-Ölçüm, G. Kiss, C. F. Rosewall, H. M. Lovick, E. L. Ingalls, J. L. Gallaher, J. Seetharaman, S. Lew, G. T. Montelione, J. F. Hunt, F. E. Michael, K. N. Houk and D. Baker, Computational design of enone-binding proteins with catalytic activity for the morita-baylis-hillman reaction, *ACS Chem. Biol.*, 2013, **8**, 749–757.
  - 185 G. Kiss, D. Röthlisberger, D. Baker and K. N. Houk, Evaluation and ranking of enzyme designs, *Protein Sci.*, 2010, **19**, 1760–1773.
  - 186 O. Khersonsky, D. Röthlisberger, A. M. Wollacott, P. Murphy, O. Dym, S. Albeck, G. Kiss, K. N. Houk, D. Baker and D. S. Tawfik, Optimization of the in-silico-designed Kemp eliminase KE70 by computational design and directed evolution, *J. Mol. Biol.*, 2011, **407**, 391–412.
  - 187 C. Zong, C. J. Wilson, T. Shen, P. Wittung-Stafshede, S. L. Mayo and P. G. Wolynes, Establishing the entatic state in folding metallated *Pseudomonas aeruginosa* azurin, *PNAS*, 2007, **104**, 3159–3164.
  - 188 B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard and D. Baker, Atomic-Level Accuracy, *Science*, 2003, **302**, 1364–1369.
  - 189 A. Lombardi, F. Pirro, O. Maglio, M. Chino and W. F. DeGrado, De novo design of four-helix bundle metalloproteins: One scaffold, diverse reactivities, *Acc. Chem. Res.*, 2019, **52**, 1148–1159.
  - 190 K. J. Jensen, De Novo Design of Proteins, *Pept. Protein Des. Biopharm. Appl.*, 2009, **101**, 207–248.
  - 191 R. Torres Martin De Rosales, M. Faiella, E. Farquhar, L. Que, C. Andreozzi, V. Pavone, O. Maglio, F. Natri and A. Lombardi, Spectroscopic and metal-binding properties of DF3: An artificial protein able to accommodate different metal ions, *J. Biol. Inorg. Chem.*, 2010, **15**, 717–728.
  - 192 S. Q. Zhang, M. Chino, L. Liu, Y. Tang, X. Hu, W. F. DeGrado and A. Lombardi,



- De Novo Design of Tetranuclear Transition Metal Clusters Stabilized by Hydrogen-Bonded Networks in Helical Bundles, *J. Am. Chem. Soc.*, 2018, **140**, 1294–1304.
- 193 S. Ramaswamy, H. Eklund and B. V Plapp, Structures of horse liver alcohol dehydrogenase complexed with NAD<sup>+</sup> and substituted benzyl alcohols, *Biochemistry*, 1994, **33**, 5230–5237.
  - 194 LigPrep, Schrödinger, LLC, New York, NY, 2016.
  - 195 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins, *J. Chem. Theory Comput.*, 2016, **12**, 281–296.
  - 196 J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, Epik: A software program for pKa prediction and protonation state generation for drug-like molecules, *J. Comput. Aided. Mol. Des.*, 2007, **21**, 681–691.
  - 197 B. V. Plapp and S. Ramaswamy, Atomic-resolution structures of horse liver alcohol dehydrogenase with NAD<sup>+</sup> and fluoroalcohols define strained michaelis complexes, *Biochemistry*, 2012, **51**, 4035–4048.
  - 198 G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments, *J. Comput. Aided. Mol. Des.*, 2013, **27**, 221–234.
  - 199 Prime, Schrodinger, LLC, New York, NY, 2016, 2002, **320**, 597–608.
  - 200 M. H. M. Olsson, C. R. SØndergaard, M. Rostkowski and J. H. Jensen, PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.
  - 201 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening, *J. Med. Chem.*, 2004, **47**, 1750–1759.
  - 202 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, *J. Med. Chem.*, 2004, **47**, 1739–1749.
  - 203 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J. Med. Chem.*, 2006, **49**, 6177–6196.
  - 204 S. Genheden and U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities., *Expert Opin. Drug Discov.*, 2015, **10**, 449–461.

- 205 J. M. Hayes and G. Archontis, in *Molecular Dynamics – Studies of Synthetic and Biological Macromolecules*, 2012, pp. 171–190.
- 206 R. A. Jianing Li, The VSGB 2.0 Model: A Next Generation Energy Model for High Resolution Protein Structure Modeling, *Proteins*, 2012, **79**, 2794–2812.
- 207 W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner and R. Farid, Novel procedure for modeling ligand/receptor induced fit effects, *J. Med. Chem.*, 2006, **49**, 534–553.
- 208 A. J. Clark, P. Tiwary, K. Borrelli, S. Feng, E. B. Miller, R. Abel, R. A. Friesner and B. J. Berne, Prediction of Protein-Ligand Binding Poses via a Combination of Induced Fit Docking and Metadynamics Simulations, *J. Chem. Theory Comput.*, 2016, **12**, 2990–2998.
- 209 H. Alonso, A. A. Bliznyuk and J. E. Gready, Combining docking and molecular dynamic simulations in drug design, *Med. Res. Rev.*, 2006, **26**, 531–568.
- 210 J. Kirchmair, P. Markt, S. Distinto, G. Wolber and T. Langer, Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection - What can we learn from earlier mistakes?, *J. Comput. Aided. Mol. Des.*, 2008, **22**, 213–228.
- 211 S. Al-Karadaghi, E. S. Cedergren-Zeppezauer, S. Hovmoller, K. Petratos, H. Terry and K. S. Wilson, Refined crystal structure of liver alcohol dehydrogenase-NADH complex at 1.8 Angstroms resolution, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1994, **50**, 793–807.
- 212 R. Meijers, R. J. Morris, H. W. Adolph, A. Merli, V. S. Lamzin and E. S. Cedergren-Zeppezauer, On the Enzymatic Activation of NADH, *J. Biol. Chem.*, 2001, **276**, 9316–9321.
- 213 J. K. Rubach and B. V. Plapp, Amino acid residues in the nicotinamide binding site contribute to catalysis by horse liver alcohol dehydrogenase., *Biochemistry*, 2003, **42**, 2907–2915.
- 214 F. Fan, J. A. Lorenzen and B. V. Plapp, An Aspartate Residue in Yeast Alcohol Dehydrogenase I Determines the Specificity for Coenzyme, *Biochemistry*, 1991, **30**, 6397–6401.
- 215 A. Johansson, K. Mosbach and M. Mansson, Horse liver alcohol dehydrogenase can accept NADP<sup>+</sup> as coenzyme in high concentrations of acetonitrile, *Eur. J. Biochem.*, 1995, **227**, 551–555.
- 216 M. Szaszko, I. Hajdú, B. Flachner, K. Dobi and Z. Kapui, Identification of potential glutamyl cyclase inhibitors from lead-like libraries by in silico and in vitro fragment-based screening, *Mol Divers*, 2017, **21**, 175–186.
- 217 H. Eklund, B. V. Plapp, J. P. Samama and C. I. Brändén, Binding of substrate in a ternary complex of horse liver alcohol dehydrogenase., *J. Biol. Chem.*, 1982, **257**, 14349–14358.
- 218 C. E. Paul, I. W. C. E. Arends and F. Hollmann, Is simpler better? Synthetic

- nicotinamide cofactor analogues for redox chemistry, *ACS Catal.*, 2014, **4**, 788–797.
- 219 J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, ZINC: A free tool to discover chemistry for biology, *J. Chem. Inf. Model.*, 2012, **52**, 1757–1768.
  - 220 H. Huo, Z. Zhou, A. Zhang and L. Wu, Ruthenium(II)-catalyzed transfer hydrogenation of aldehydes with new water-soluble monotosylated ethylenediamines as ligands, *Res. Chem. Intermed.*, 2012, **38**, 261–268.
  - 221 K. Krohn, H. Dörner and M. Zukowski, Chemical synthesis of benzamide riboside., *Curr. Med. Chem.*, 2002, **9**, 727–731.
  - 222 A. M. Martín Castro, Claisen Rearrangement over the Past Nine Decades, *Chem. Rev.*, 2004, **104**, 2939–3002.
  - 223 M. S. Abdel-Maksoud, M.-R. Kim, M. I. El-Gamal, M. M. Gamal El-Din, J. Tae, H. S. Choi, K.-T. Lee, K. H. Yoo and C.-H. Oh, Design, synthesis, in vitro antiproliferative evaluation, and kinase inhibitory effects of a new series of imidazo[2,1-b]thiazole derivatives, *Eur. J. Med. Chem.*, 2015, **95**, 453–463.
  - 224 A. Dijksman, J. M. Elzinga, Y. X. Li, I. W. C. E. Arends and R. A. Sheldon, Efficient ruthenium-catalyzed racemization of secondary alcohols: Application to dynamic kinetic resolution, *Tetrahedron Asymmetry*, 2002, **13**, 879–884.
  - 225 J. G. Kim and D. O. Jang, A convenient one-pot method for the synthesis of N-methoxy-N-methyl amides from carboxylic acids, *Bull. Korean Chem. Soc.*, 2010, **31**, 171–173.
  - 226 Y. Tahara, M. Ito, K. S. Kanyiva and T. Shibata, Total Synthesis of cis -Clavicipitic Acid from Asparagine via Ir-Catalyzed C-H bond Activation as a Key Step, *Chem. A Eur. J.*, 2015, **21**, 11340–11343.
  - 227 E. Csimbók, D. Takács, J. A. Balog, O. Egyed, N. V. May-Nagy and G. M. Keserű, The first synthesis of isoxazolo[3,4-c]pyridine-7-ones, *Tetrahedron Lett.*, 2016, **57**, 4401–4404.
  - 228 B. Chen, H. F. Yin, Z. S. Wang and J. H. Xu, New synthesis of harzialactone A via kinetic resolution using recombinant *Fusarium proliferatum* lactonase, *Tetrahedron Asymmetry*, 2010, **21**, 237–240.
  - 229 L. Brecker, M. Pogorevc, H. Griengl, W. Steiner, T. Kappe and D. W. Ribbons, Synthesis of 2,4-diketoacids and their aqueous solution structures, *New J. Chem.*, 1999, **23**, 437–446.
  - 230 Y. Zhao, Q. Wang, Q. Meng, D. Ding, H. Yang, G. Gao, D. Li, W. Zhu and H. Zhou, Identification of Trypanosoma brucei leucyl-tRNA synthetase inhibitors by pharmacophore- and docking-based virtual screening and synthesis, *Bioorganic Med. Chem.*, 2012, **20**, 1240–1250.
  - 231 F. W. Swamer and C. R. Hauser, Claisen Acylations and Carbethoxylations of Ketones and Esters by Means of Sodium Hydride, *J. Am. Chem. Soc.*, 1950, **72**,

1352–1356.

- 232 M. K. Wong, J. L. Ali, SYNTHESIS OF  $\alpha$ -ARYL- $\beta$ -KEMESTER, *Org. Prep. Proced.*, 1970, **2**, 193–195.
- 233 J. A. Jiang, W. Bin Huang, J. J. Zhai, H. W. Liu, Q. Cai, L. X. Xu, W. Wang and Y. F. Ji, 'One-pot' synthesis of 4-substituted 1,5-diaryl-1H-pyrazole-3-carboxylates via lithium tert-butoxide-mediated sterically hindered Claisen condensation and Knorr reaction, *Tetrahedron*, 2013, **69**, 627–635.
- 234 M. M. Schiavoni, H. E. Di Loreto, A. Hermann, H. Mack, S. E. Ulic and C. O. Della, Keto – enol tautomerism in  $\beta$ -ketoesters:  $\text{CH}_3\text{C}(\text{O})\text{CHX}\text{C}(\text{O})\text{OY}$  ( $\text{X} = \text{H}, \text{Cl}; \text{Y} = \text{CH}_3, \text{C}_2\text{H}_5$ ). Vibrational analyses, NMR spectra and quantum chemical calculations, *J. Raman Spectrosc.*, 2001, **32**, 319–329.
- 235 A. I. Kol'tsov and G. M. Kheifets, Investigation of Keto–Enol Tautomerism by Nuclear Magnetic Resonance Spectroscopy, *Russ. Chem. Rev.*, 1971, **40**, 773–788.
- 236 T. Rahn, T. H. T. Dang, A. Spannenberg, C. Fischer and P. Langer, Regioselective synthesis of functionalized 3,5-diketoesters and 2,4-diketosulfones by uncatalyzed condensation of 1-methoxy-1,3-bis(trimethylsilyloxy)-1,3-butadienes with  $\alpha,\beta$ -unsaturated acid chlorides and sulfonyl chlorides, *Org. Biomol. Chem.*, 2008, **6**, 3366–3370.
- 237 P. G. Greene, T. W.; Wuts, in *4rd Edition, John Wiley, New York*, 2007, pp. 1794–1798.
- 238 C. G. Boojamra, C. Cannizzaro, J. M. Chen, X. Chen, A. Cho, L. S. Chong, M. Fardis and H. Jin, Antiviral phosphonate analogs, 2008, **2**, 1–393.
- 239 X. Wang, L. Xu, F. Xiong, Y. Wu and F. Chen, A new cost-effective Ru-chloramphenicol base derivative catalyst for the asymmetric transfer hydrogenation/dynamic kinetic resolution of N-Boc  $\alpha$ -amino- $\beta$ -ketoesters and its application to the synthesis of the chiral core of vancomycin, *RSC Adv.*, 2016, **6**, 37701–37709.
- 240 A. R. Mohite, P. R. Sultane and R. G. Bhat,  $\text{BF}_3\cdot\text{Et}_2\text{O}$  and trifluoroacetic acid/triethyl amine-mediated synthesis of functionalized piperidines, *Tetrahedron Lett.*, 2012, **53**, 30–35.
- 241 R. P. Pandit and Y. R. Lee, Novel one-pot synthesis of diverse  $\gamma,\delta$ -unsaturated  $\beta$ -ketoesters by thermal cascade reactions of diazodicarbonyl compounds and enol ethers: Transformation into substituted 3,5-diketoesters, *Org. Biomol. Chem.*, 2014, **12**, 4407–4411.
- 242 M. M. Faul, T. A. Engler, K. A. Sullivan, J. L. Grutsch, M. T. Clayton, M. J. Martinelli, J. M. Pawlak, M. LeTourneau, D. S. Coffey, S. W. Pedersen, S. P. Kolis, K. Furness, S. Malhotra, R. S. Al-Awar and J. E. Ray, Synthetic Approaches to Indolo[6,7- $\alpha$ ]pyrrolo[3,4-c]carbazoles: Potent Cyclin D1/CDK4 Inhibitors, *J. Org. Chem.*, 2004, **69**, 2967–2975.
- 243 P. S. Da Silva, T. Rossi, L. E. Kiss, A. S. Beliaev and P. N. L. Palma, Dopamine-b-

hydroxylase inhibitors field, 2020, 1–30.

- 244 S. Gibson, S. C. Bergmeier and H. Rapoport, Selective Removal of an N-BOC Protecting Group in the Presence of a tert-Butyl Ester and Other Acid-Sensitive Groups, *J. Org. Chem.*, 1994, **59**, 3216–3218.
- 245 R. A. Houghten, A. Beckman and J. M. Ostresh, Use of 10% sulfuric acid/dioxane for removal of N- $\alpha$ -tertiary-butyloxycarbonyl group during solid phase peptide synthesis, *Int. J. Pept. Protein Res.*, 1986, **27**, 653–658.
- 246 P. Strazzolini, N. Misuri and P. Polese, Efficient cleavage of carboxylic tert-butyl and 1-adamantyl esters, and N-Boc-amines using H<sub>2</sub>SO<sub>4</sub> in CH<sub>2</sub>Cl<sub>2</sub>, *Tetrahedron Lett.*, 2005, **46**, 2075–2078.
- 247 C. Zinelaabidine, O. Souad, J. Zoubir, B. Malika and A. Nour-Eddine, A Simple and Efficient Green Method for the Deprotection of N-Boc in Various Structurally Diverse Amines under Water-mediated Catalyst-free Conditions, *Int. J. Chem.*, 2012, **4**, 73–79.
- 248 G. Wang, C. Li, J. Li and X. Jia, Catalyst-free water-mediated N -Boc deprotection, *Tetrahedron Lett.*, 2009, **50**, 1438–1440.
- 249 G. Pavan Kumar, D. Rambabu, M. V. Basaveswara Rao and M. Pal, Iodine-mediated neutral and selective N-boc deprotection, *J. Chem.*, 2013, **2013**, 1–5.
- 250 K. Krohn, A. Vidal, J. Vitz, B. Westermann, M. Abbas and I. Green, First enantiospecific Baker-Venkataraman-rearrangements aiming at the total synthesis of chiral anthracycline antibiotics, *Tetrahedron Asymmetry*, 2006, **17**, 3051–3057.
- 251 R. Sundaramoorthi, C. Siedem, C. B. Vu, D. C. Dalgarno, E. C. Laird, M. C. Botfield, A. B. Combs, S. E. Adams, R. W. Yuan, M. Weigle and S. S. Narula, Selective inhibition of Src SH2 by a novel thiol-targeting tricarbonyl-modified inhibitor and mechanistic analysis by <sup>1</sup>H/ <sup>13</sup>C NMR spectroscopy, *Bioorganic Med. Chem. Lett.*, 2001, **11**, 1665–1669.
- 252 S. G. Johnson, P. J. Connolly and W. V. Murray, A concise synthesis of 7-chloro-2-methylsulfanyl-thiazolo[4,5-b]pyridine-6-carbonitrile, a versatile intermediate for substituted 6-cyanothiazolopyridines, *Tetrahedron Lett.*, 2006, **47**, 4853–4856.
- 253 G. Lang, A. L. J. Cole, J. W. Blunt and M. H. G. Munro, An unusual oxalylated tetramic acid from the New Zealand basidiomycete *Chamonixia pachydermis*, *J. Nat. Prod.*, 2006, **69**, 151–153.
- 254 N. H. Pungot, Z. Shaameri, A. S. Hamzah, M. F. Mohammat and N. Hussain, Synthesis of 3-methyl-5-nitrobenzyl  $\beta$ ,  $\beta$ -diketoester as a derivative of pachydermin, a tetramic acid from *Chamonixia pachydermis*, *J. Sustain. Sci. Manag.*, 2017, **12**, 1–7.
- 255 R. Braga, L. Hecquet and C. Blonski, Slow-binding inhibition of 2-keto-3-deoxy-6-phosphogluconate ( KDPG ) aldolase, *Bioorg. Med. Chem.*, 2004, **12**, 2965–

2972.

- 256 M. Jukic, D. Sterk and Z. Casar, Recent Advances in the Retro-Claisen Reaction and Its Synthetic Applications, *Curr. Org. Synth.*, 2012, **9**, 488–512.
- 257 N. W. Fadnavis and K. R. Radhika, Enantio- and regiospecific reduction of ethyl 4-phenyl-2,4-dioxobutyrate with baker's yeast: Preparation of (R)-HPB ester, *Tetrahedron Asymmetry*, 2004, **15**, 3443–3447.
- 258 Y. L. Song, M. L. Peach, P. P. Roller, S. Qiu, S. Wang and Y. Q. Long, Discovery of a novel nonphosphorylated pentapeptide motif displaying high affinity for Grb2-SH2 domain by the utilization of 3'-substituted tyrosine derivatives, *J. Med. Chem.*, 2006, **49**, 1585–1596.
- 259 R. A. Croft, M. A. J. Dubois, A. J. Boddy, C. Denis, A. Lazaridou, A. S. Voisin-chiret, R. Bureau, C. Choi, J. J. Mousseau and J. A. Bull, Catalytic Friedel-Crafts Reactions on Saturated Medicinally Relevant Fragments, *European J. Org. Chem.*, 2019, **2019**, 5385–5395.
- 260 J. M. Sprague and H. Adkins, Hydrogenation and Hydrogenolysis of 1,3-Diketones, *J. Am. Chem. Soc.*, 1934, **56**, 2669–2675.
- 261 L. F. Kuick and H. Adkins, The Preparation, Alcoholysis and Hydrogenolysis of Nicotiny Acyl Methanes, *J. Am. Chem. Soc.*, 1935, **57**, 143–147.
- 262 C. M. Osmundsen, K. Egeblad and E. Taarning, in *New and Future Developments in Catalysis: Catalytic Biomass Conversion*, Elsevier B.V., 2013, pp. 73–89.
- 263 M. V. Galkin, C. Dahlstrand and J. S. M. Samec, Mild and Robust Redox-Neutral Pd/C-Catalyzed Lignol  $\beta$ -O-4' Bond Cleavage Through a Low-Energy-Barrier Pathway, *ChemSusChem*, 2015, **8**, 2187–2192.
- 264 M. Studer, S. Burkhardt, A. F. Indolese and H. U. Blaser, Enantio- and chemoselective reduction of 2,4-diketo acid derivatives with cinchona modified Pt-catalyst - Synthesis of (R)-2-hydroxy-4- phenylbutyric acid ethyl ester, *Chem. Commun.*, 2000, 1327–1328.
- 265 P. Herold, A. F. Indolese, M. Studer, H. P. Jalett, U. Siegrist and H. U. Blaser, New technical synthesis of ethyl (R)-2-hydroxy-4-phenylbutyrate of high enantiomeric purity, *Tetrahedron*, 2000, **56**, 6497–6499.
- 266 S. Nishimura, *Handbook of heterogenous catalytic hydrogenation for organic synthesis*, 1962, vol. 54.
- 267 B. Chen, E. R. Jamieson and T. D. Tullius, A general synthesis of specifically deuterated nucleotides for studies of DNA and RNA, *Bioorganic Med. Chem. Lett.*, 2002, **12**, 3093–3096.
- 268 W. Xie, G. Tanabe, K. Matsuoka, M. F. A. Amer, T. Minematsu, X. Wu, M. Yoshikawa and O. Muraoka, Role of the side chain stereochemistry in the  $\alpha$ -glucosidase inhibitory activity of kotalanol, a potent natural  $\alpha$ -glucosidase inhibitor, *Bioorganic Med. Chem.*, 2011, **19**, 2252–2262.

- 269 I. R. Corrêa and P. J. S. Moran, Diastereoselective reduction of E and Z  $\alpha$ -alkoxyimino- $\beta$ -ketoesters by sodium borohydride, *Tetrahedron*, 1999, **55**, 14221–14232.
- 270 K. Sivagurunathan, S. Raja Mohamed Kamil, S. Syed Shafi, F. Liakth Ali Khan and R. V. Ragavan, Efficient one-pot selective reduction of esters in  $\beta$ -ketoesters using LiHMDS and lithium aluminium hydride, *Tetrahedron Lett.*, 2011, **52**, 1205–1207.
- 271 S. Veeraswamy, K. Indrasena Reddy, R. Venkat Ragavan, K. Tirumal Reddy, S. Yennam and A. Jayashree, An efficient one-step chemoselective reduction of alkyl ketones over aryl ketones in  $\beta$ -diketones using LiHMDS and lithium aluminium hydride, *Tetrahedron Lett.*, 2012, **53**, 4651–4653.
- 272 L. Knorr, Synthese von Furfuranderivaten aus dem Diacetbernsteinsäureester, *Berichte der Dtsch. Chem. Gesellschaft*, 1884, **17**, 2863–2870.
- 273 C. Paal, Ueber die Derivate des Acetophenonacetessigesters und des Acetylacetessigesters, *Berichte der Dtsch. Chem. Gesellschaft*, 1884, **17**, 2756–2767.
- 274 V. Amarnath, D. C. Anthony, K. Amarnath, W. M. Valentine, L. A. Wetterau and D. G. Graham, Intermediates in the Paal-Knorr Synthesis of Pyrroles, *J. Org. Chem.*, 1991, **56**, 6924–6931.
- 275 G. Szakál-Quin, D. G. Graham, D. S. Millington, D. A. Maltby and A. T. McPhail, Stereoisomer Effects on the Paal-Knorr Synthesis of Pyrroles, *J. Org. Chem.*, 1986, **51**, 621–624.
- 276 V. Amarnath and K. Amarnath, Intermediates in the Paal-Knorr Synthesis of Furans, *J. Org. Chem.*, 1995, **60**, 301–307.
- 277 S. Abbat, D. Dhaked, M. Arfeen and P. V. Bharatam, Mechanism of the Paal-Knorr reaction: The importance of water mediated hemiacetal pathway, *RSC Adv.*, 2015, **5**, 88353–88366.
- 278 L. D. Krasnoslobodskaya and Y. L. Gol'dfarb, Methods of Synthesis of  $\beta$ -Substituted Furan Derivatives, *Russ. Chem. Rev.*, 1969, **38**, 389–406.
- 279 B. C. L. Heilbron, I. M. Jones, E. R. H. Smith, P. Weedon, Researches on Acetylenic Compounds. Part IV. The Hydration of Xome Acetylenic Compounds Derived From up- Unsaturated Aldehydes., *J. Chem. Soc.*, 1946, **54**, 54–58.
- 280 B. C. L. Bowden, K., Braude, E. A. Jones, E. R. H. Weedon, Researches on Acetylenic Compounds. Part II. (A) The Addition of Amines to Ethynyl Ketones. (B) Auxochromic Properties and Conjugating Power of the Amino Group., *J. Chem. Soc.*, 1946, **45**, 45–52.
- 281 F. Stauffer and R. Neier, Synthesis of tri- and tetrasubstituted furans catalyzed by trifluoroacetic acid, *Org. Lett.*, 2000, **2**, 3535–3537.
- 282 J. Corpas, A. Ponce, J. Adrio and J. C. Carretero, CuI-Catalyzed Asymmetric [3 + 2] Cycloaddition of Azomethine Ylides with Cyclobutenones, *Org. Lett.*, 2018,

**20**, 3179–3182.

- 283 A. A. Taubinger, D. Fenske and J. Podlech, Synthesis of  $\beta,\beta'$ -diamino acids from  $\alpha$ -amino acid-derived  $\beta$ -lactams by ring opening with nucleophiles. Utilization in the synthesis of peptidomimetics, *Tetrahedron*, 2008, **64**, 8659–8667.
- 284 H. van Bekkum, A. P. G. Kieboom and K. J. G. van De Putte, Electronic and steric effects in the hydrogenation of alkyl aryl ketones on palladium, *Recl. des Trav. Chim. des Pays-Bas*, 1969, **88**, 52–61.
- 285 J. W. Grate, K. Mo and M. D. Daily, Triazine-Based Sequence-Defined Polymers with Side-Chain Diversity and Backbone–Backbone Interaction Motifs, *Angew. Chemie*, 2016, **128**, 3993–3998.
- 286 D. W. Lee, H. J. Ha and W. K. Lee, Selective mono-BOC protection of diamines, *Synth. Commun.*, 2007, **37**, 737–742.
- 287 A. M. Nethercott, Engineering of a Thermostable Alcohol Dehydrogenase Towards Imine Reduction, *Univ. Nottingham*, 2017, 1–206.
- 288 D. Quaglia, J. A. Irwin and F. Paradisi, Horse liver alcohol dehydrogenase: New perspectives for an old enzyme, *Mol. Biotechnol.*, 2012, **52**, 244–250.
- 289 S. Morra and A. Pordea, Biocatalyst-artificial metalloenzyme cascade based on alcohol dehydrogenase, *Chem. Sci.*, 2018, **9**, 7447–7454.
- 290 M. Peretz, O. Bogin, S. Tel-Or, A. Cohen, G. Li, J. S. Chen and Y. Burstein, Molecular cloning, nucleotide sequencing, and expression of genes encoding alcohol dehydrogenases from the thermophile *Thermoanaerobacter brockii* and the mesophile *Clostridium beijerinckii*, *Anaerobe*, 1997, **3**, 259–270.
- 291 Y. Okamoto, V. Köhler and T. R. Ward, An NAD ( P ) H-dependent Artificial Transfer Hydrogenase for Multi-enzymatic, *J. Am. Chem. Soc.*, 2016, **138**, 5781–5784.
- 292 S. Keller, C. Vargas, H. Zhao, G. Piszczek, C. A. Brautigam and P. Schuck, High-precision isothermal titration calorimetry with automated peak-shape analysis, *Anal. Chem.*, 2012, **84**, 5066–5073.
- 293 Z. R. Marjenberg, I. R. Ellis, R. M. Hagan, S. Prabhakaran, M. Höök, S. R. Talay, J. R. Potts, D. Staunton and U. Schwarz-Linek, Cooperative binding and activation of fibronectin by a bacterial surface protein, *J. Biol. Chem.*, 2011, **286**, 1884–1894.
- 294 K. H. Dahl and J. S. McKinley-McKee, The Imidazole-Promoted Inactivation of Horse-Liver Alcohol Dehydrogenase, *Eur. J. Biochem.*, 1981, **120**, 451–459.
- 295 R. J. Lamed and J. G. Zeikus, Novel NADP-linked alcohol-aldehyde/ketone oxidoreductase in thermophilic ethanologenic bacteria, *Biochem. J.*, 1981, **195**, 183–190.
- 296 J. Broecker, C. Vargas and S. Keller, Revisiting the optimal c value for isothermal titration calorimetry, *Anal. Biochem.*, 2011, **418**, 307–309.



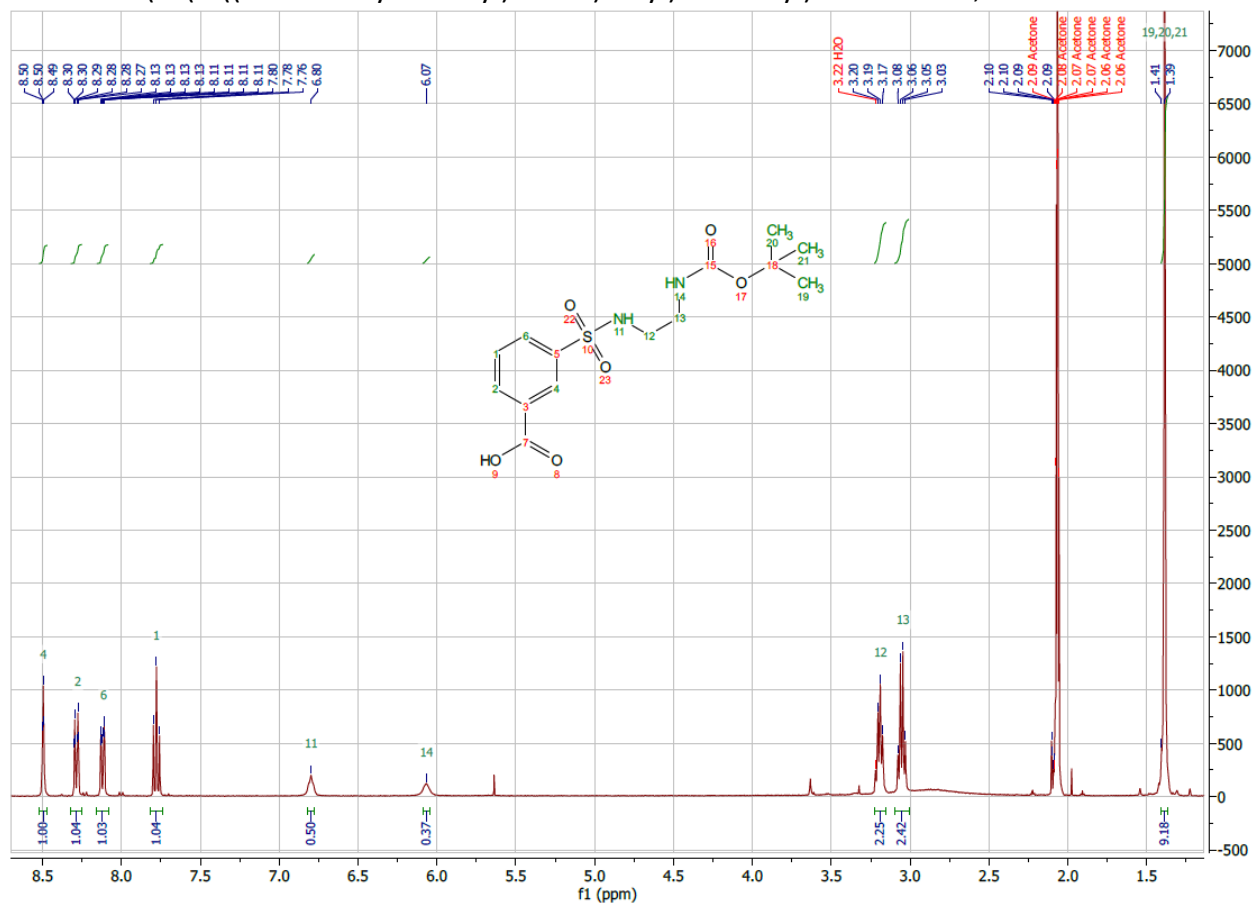
- 297 TA Instruments, *MCAPN-2016-1*, 2016, 1–5.
- 298 P. D. Subramanian, S. Ross, Calorimetry investigation of NAD binding to some dehydrogenases, *Biochem. Biophys. Res. Commun.*, 1977, **78**, 461–466.
- 299 S. Subramanian and P. D. Ross, Thermodynamics of Binding of Oxidized and Reduced Nicotinamide Adenine Dinucleotides, Adenosine-5'-diphosphoribose, and 5'-Iodosalicylate to Dehydrogenases, *Biochemistry*, 1978, **17**, 2193–2197.
- 300 M. R. Kasimova, J. Grigiene, K. Krab, P. H. Hagedorn, H. Flyvbjerg, P. E. Andersen and I. M. Møller, The free NADH concentration is kept constant in plant mitochondria under different metabolic conditions, *Plant Cell*, 2006, **18**, 688–698.
- 301 M. Poizat, I. W. C. E. Arends and F. Hollmann, On the nature of mutual inactivation between [Cp\*Rh(bpy)(H<sub>2</sub>O)]<sup>2+</sup> and enzymes - analysis and potential remedies, *J. Mol. Catal. B Enzym.*, 2010, **63**, 149–156.
- 302 K. Terpe, Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems, *Appl. Microbiol. Biotechnol.*, 2003, **60**, 523–533.
- 303 H. A. H. Abd Elhameed, B. Hajdu, A. Jancsó, A. Kéri, G. Galbács, É. Hunyadi-Gulyás and B. Gyurcsik, Modulation of the catalytic activity of a metallonuclease by tagging with oligohistidine, *J. Inorg. Biochem.*, 2020, **206**, 1–11.
- 304 J. Watly, E. Simonovsky, R. Wieczorek, N. Barbosa, Y. Miller and H. Kozłowski, Insight into the coordination and the binding sites of Cu<sup>2+</sup> by the histidyl-6-tag using experimental and computational tools, *Inorg. Chem.*, 2014, **53**, 6675–6683.
- 305 J. Wąty, A. Hecel, M. Rowińska-Żyrek and H. Kozłowski, Impact of histidine spacing on modified polyhistidine tag – Metal ion interactions, *Inorganica Chim. Acta*, 2018, **472**, 119–126.
- 306 D. Brasili, J. Watly, E. Simonovsky, R. Guerrini, N. A. Barbosa, R. Wieczorek, M. Remelli, H. Kozłowski and Y. Miller, The unusual metal ion binding ability of histidyl tags and their mutated derivatives, *Dalt. Trans.*, 2016, **45**, 5629–5639.
- 307 V. D. Westerhausen, S. Herrmann, W. Hummel and E. Steckhan, Formate-Driven, Non-Enzymatic NAD(P)H Regeneration for the Alcohol Dehydrogenase Catalyzed Stereoselective Reduction of 4-Phenyl-2-butanone, *Angew. Chemie Int. Ed. English*, 1992, **31**, 1529–1531.
- 308 V. Muñoz Robles, E. Ortega-Carrasco, L. Alonso-Cotchico, J. Rodriguez-Guerra, A. Lledós and J. D. Maréchal, Toward the computational design of artificial metalloenzymes: From protein-ligand docking to multiscale approaches, *ACS Catal.*, 2015, **5**, 2469–2480.
- 309 M. Bastos and A. Velazquez-Campoy, Isothermal titration calorimetry (ITC): a standard operating procedure (SOP), *Eur. Biophys. J.*, 2021, 1–9.

- 310 G. Krainer and S. Keller, Single-experiment displacement assay for quantifying high-affinity binding by isothermal titration calorimetry, *Methods*, 2015, **76**, 116–123.
- 311 L. Panella, J. Broos, J. Jin, M. W. Fraaije, D. B. Janssen, M. Jeronimus-Stratingh, B. L. Feringa, A. J. Minnaard and J. G. De Vries, Merging homogeneous catalysis with biocatalysis; papain as hydrogenation catalyst, *Chem. Commun.*, 2005, **23**, 5656–5658.
- 312 V. M. Robles, M. Dürrenberger, T. Heinisch, A. Lledós, T. Schirmer, T. R. Ward and J. D. Maréchal, Structural, kinetic, and docking studies of artificial imine reductases based on biotin-streptavidin technology: An induced lock-and-key hypothesis, *J. Am. Chem. Soc.*, 2014, **136**, 15676–15683.
- 313 D. L. Ma, W. L. Wong, W. H. Chung, F. Y. Chan, P. K. So, T. S. Lai, Z. Y. Zhou, Y. C. Leung and K. Y. Wong, A highly selective luminescent switch-on probe for histidine/histidine-rich proteins and its application in protein staining, *Angew. Chemie Int. Ed.*, 2008, **47**, 3735–3739.
- 314 M. Zhou and Y. Xia, Electrochemiluminescence immunoassay based on site-specific labeling using a histidine-binding iridium(III) solvato complex, *Anal. Chim. Acta*, 2018, **1023**, 29–34.
- 315 C. Nitsche and G. Otting, NMR studies of ligand binding, *Curr. Opin. Struct. Biol.*, 2018, **48**, 16–22.

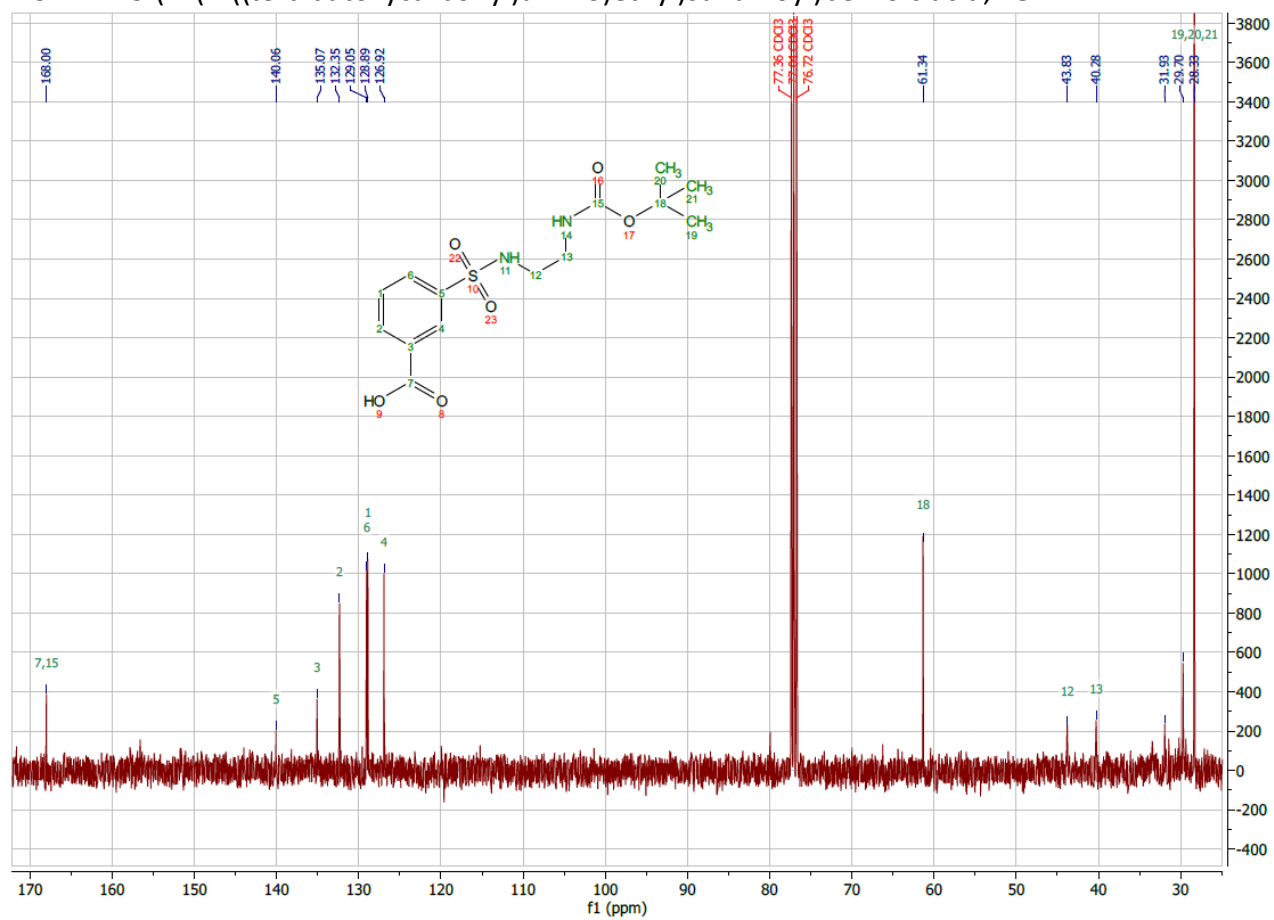
## 8 Appendices

### Appendix 1 – NMRs and ESI-MS analysis from the synthesis of a transition metal catalyst

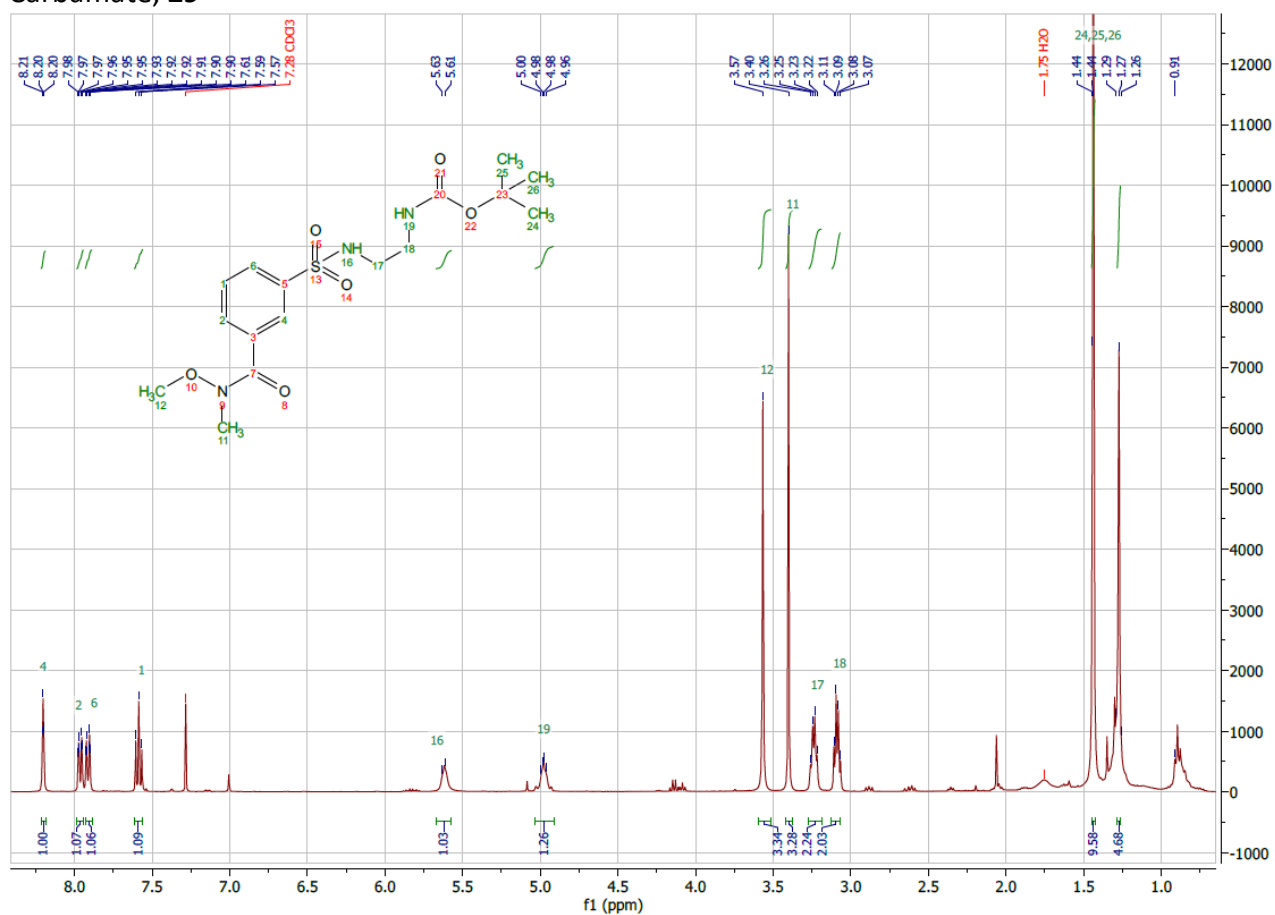
$^1\text{H}$ -NMR 3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfamoyl)benzoic acid, **28**



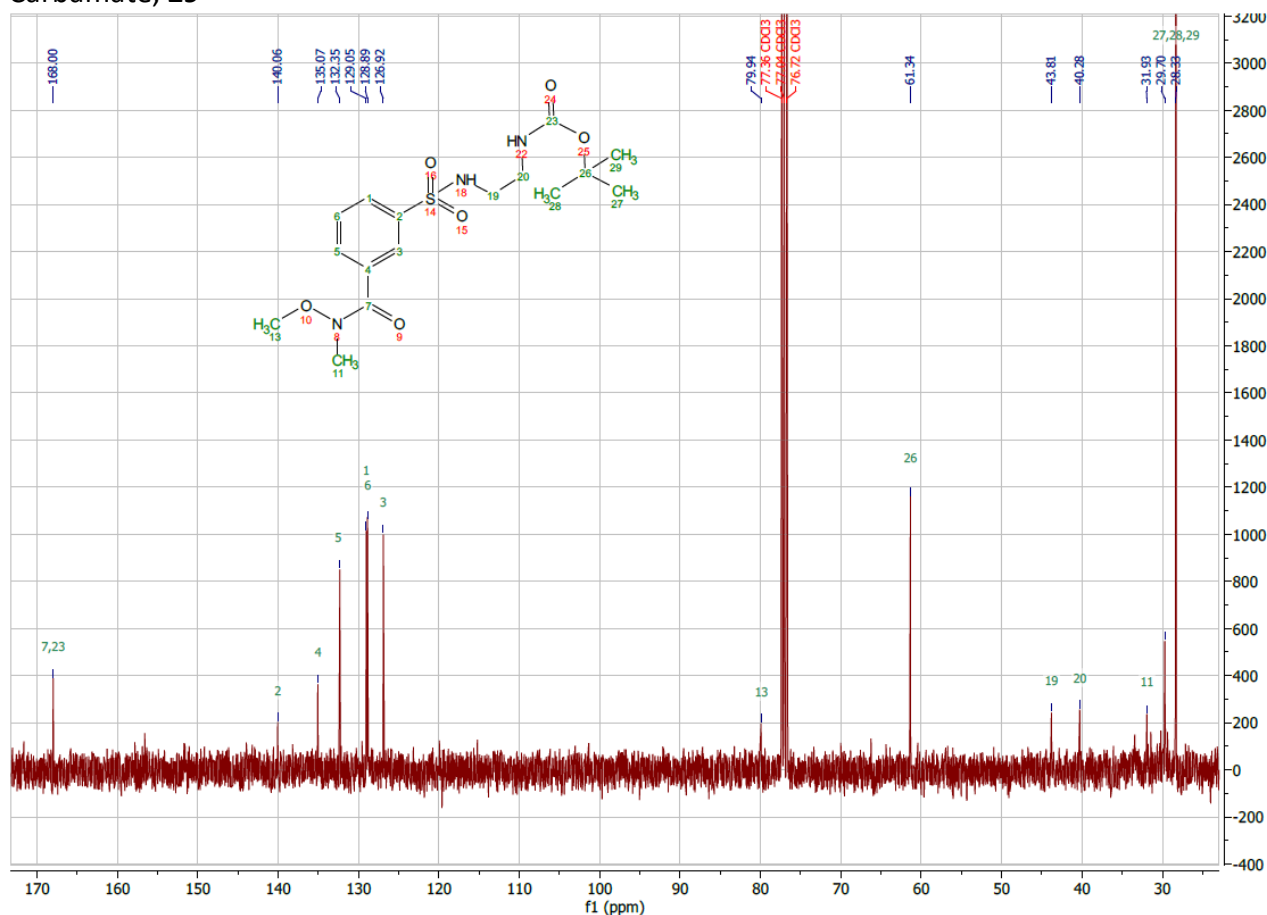
<sup>13</sup>C-NMR 3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfamoyl)benzoic acid, **28**



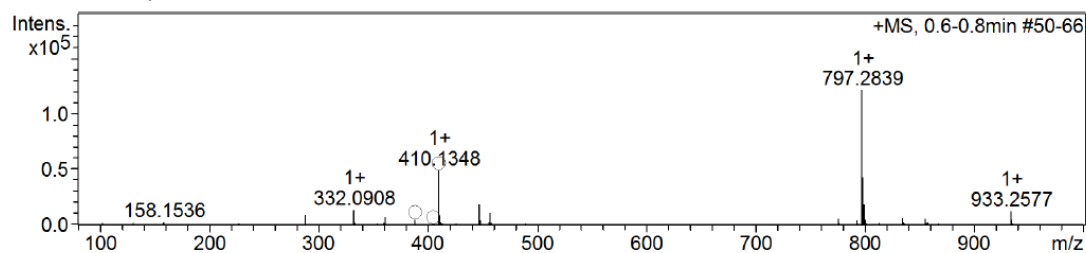
<sup>1</sup>H-NMR *tert*-butyl(2-((3(methoxy(methyl)carbamoyl)phenyl)sulfonamido)ethyl) Carbamate, **29**



<sup>13</sup>C-NMR *tert*-butyl(2-((3(methoxy(methyl)carbamoyl)phenyl)sulfonamido)ethyl) Carbamate, **29**

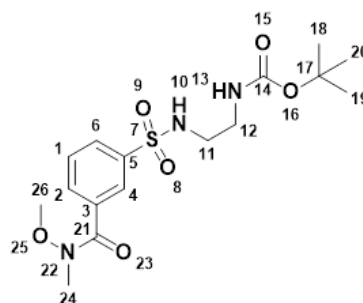


ESI-MS *tert*-butyl(2-((3(methoxy(methyl)carbamoyl)phenyl)sulfonamido)ethyl) Carbamate, **29**

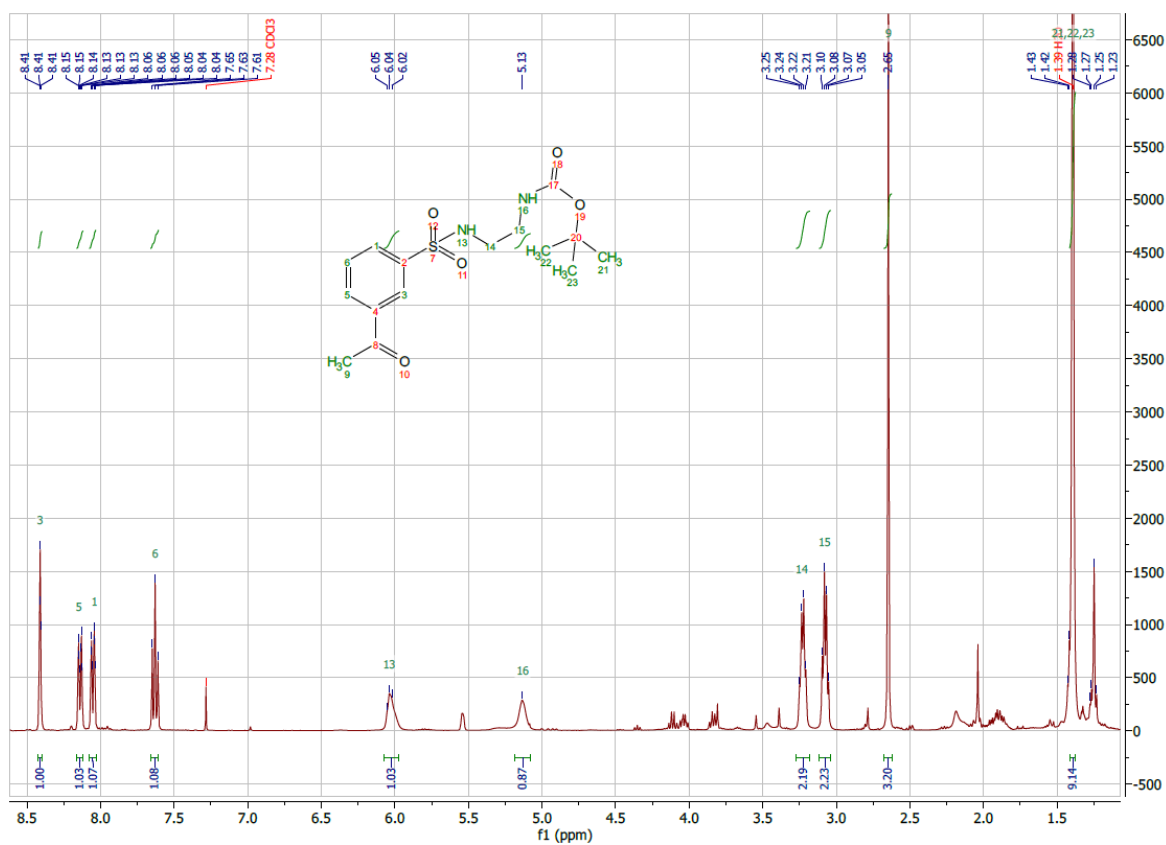


#	m/z	I %
1	288.1002	7.8
2	332.0908	11.3
3	361.0370	5.9
4	388.1545	4.2
5	410.1348	40.6
6	411.1376	7.4
7	447.2120	15.4
8	448.2120	3.4
9	457.1008	9.0
10	775.2992	5.0
11	792.3240	3.4
12	797.2839	100.0
13	798.2845	35.3
14	799.2818	15.5
15	800.2871	4.2
16	834.3510	5.4
17	835.3506	2.4
18	855.2392	5.1
19	933.2577	10.6
20	934.2565	3.9

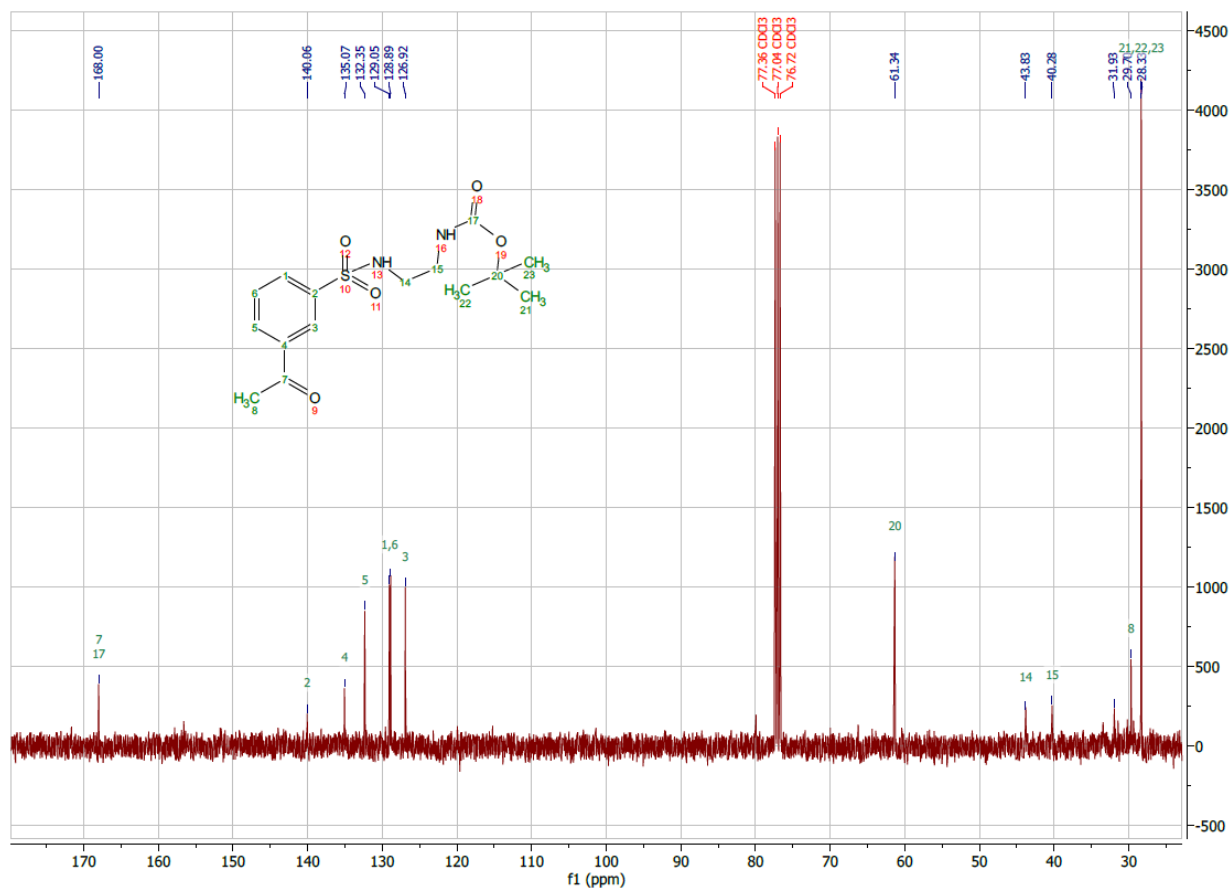
← [2M + Na<sup>+</sup>]



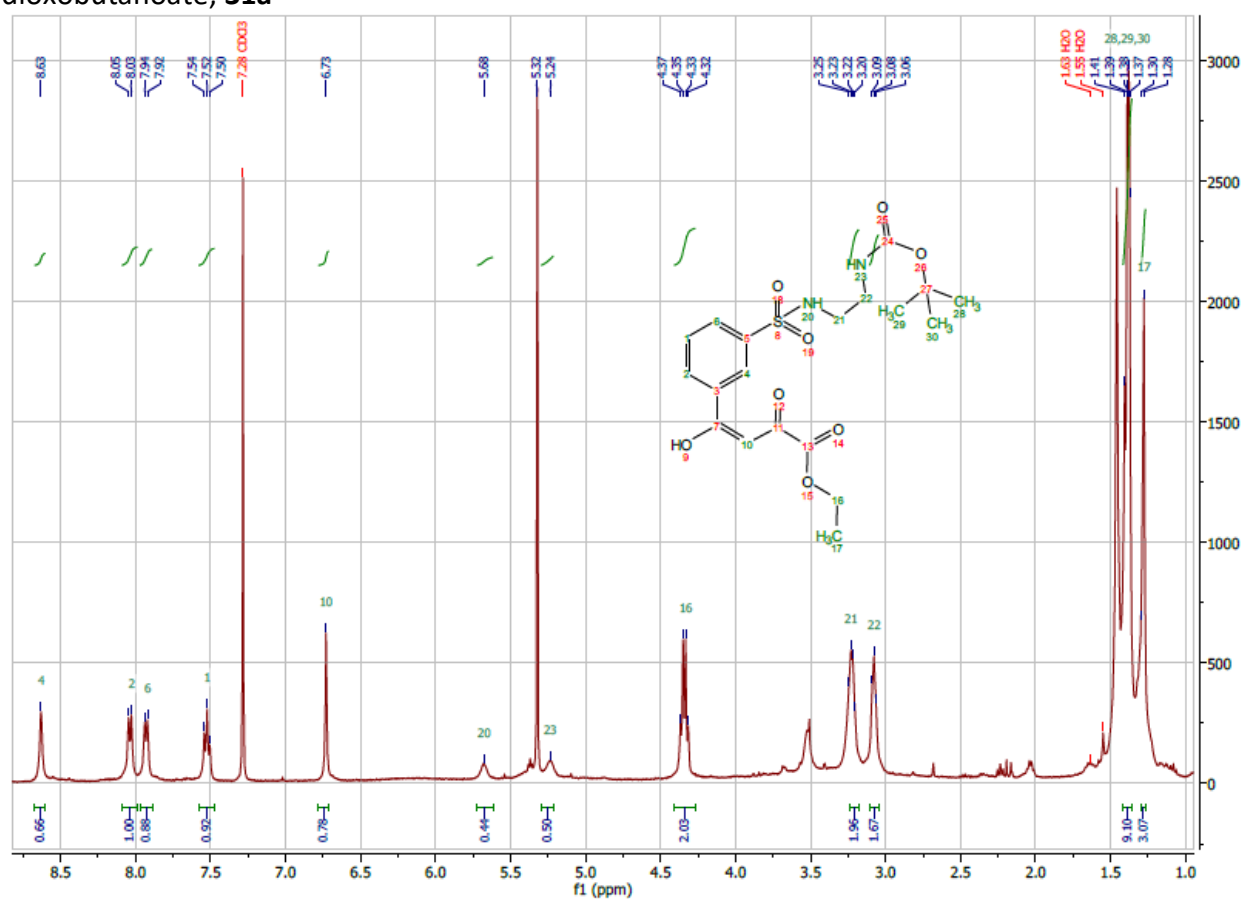
<sup>1</sup>H-NMR *tert*-butyl (2-((3-acetylphenyl)sulfonamido)ethyl)carbamate, **30**



<sup>13</sup>C-NMR *tert*-butyl (2-((3-acetylphenyl)sulfonamido)ethyl)carbamate, **30**

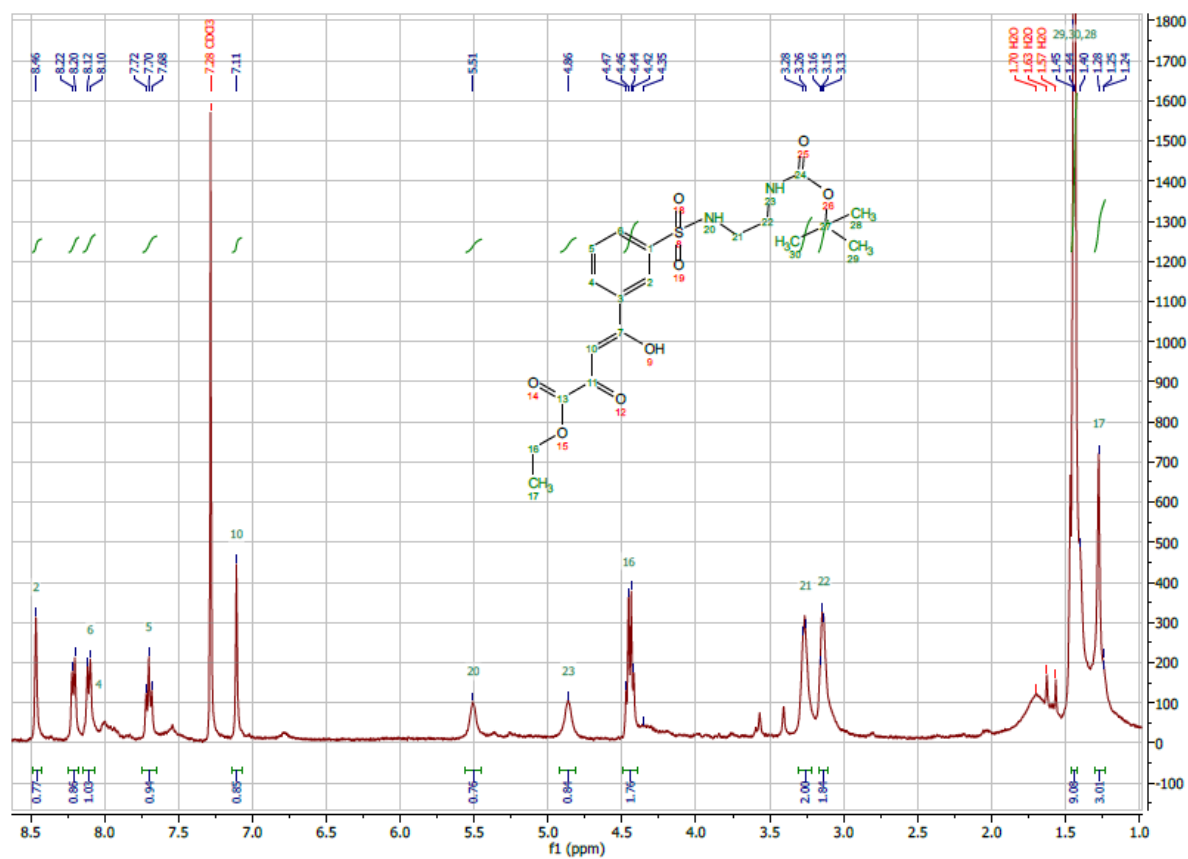


<sup>1</sup>H-NMR Ethyl 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **31a**

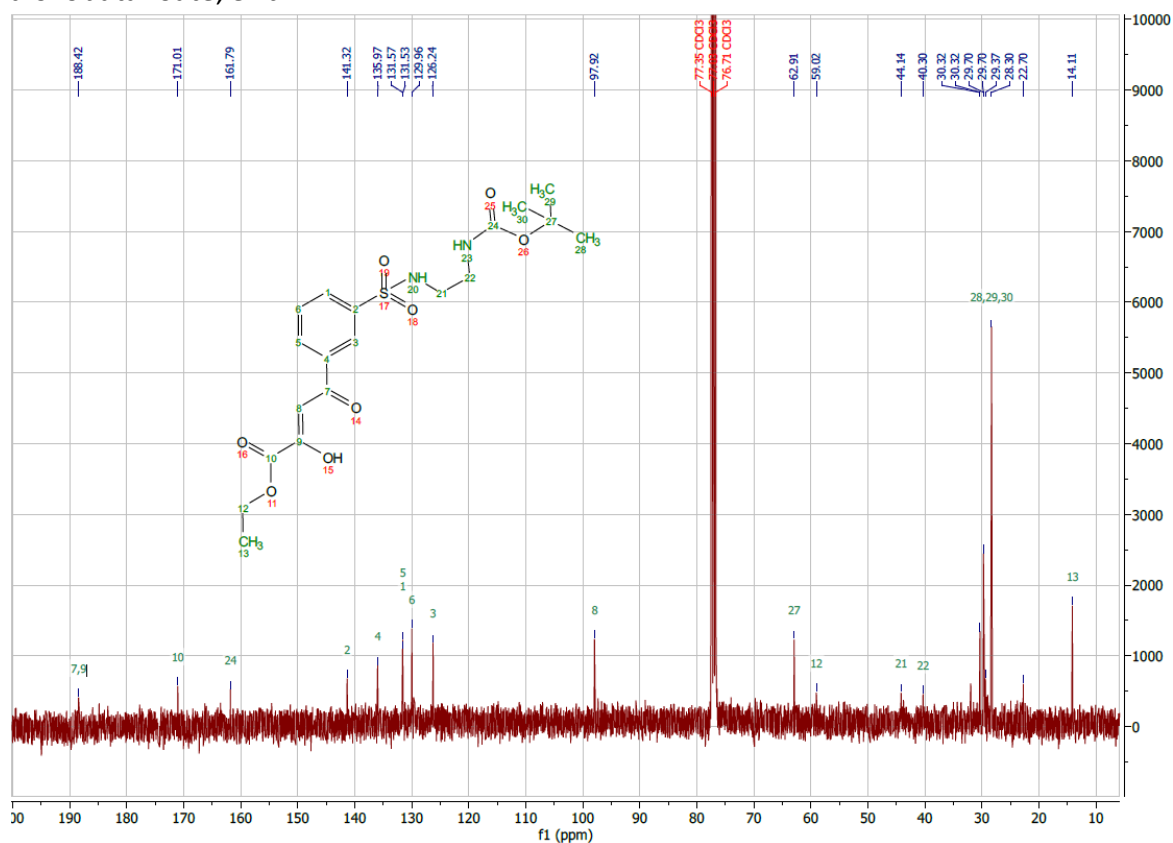




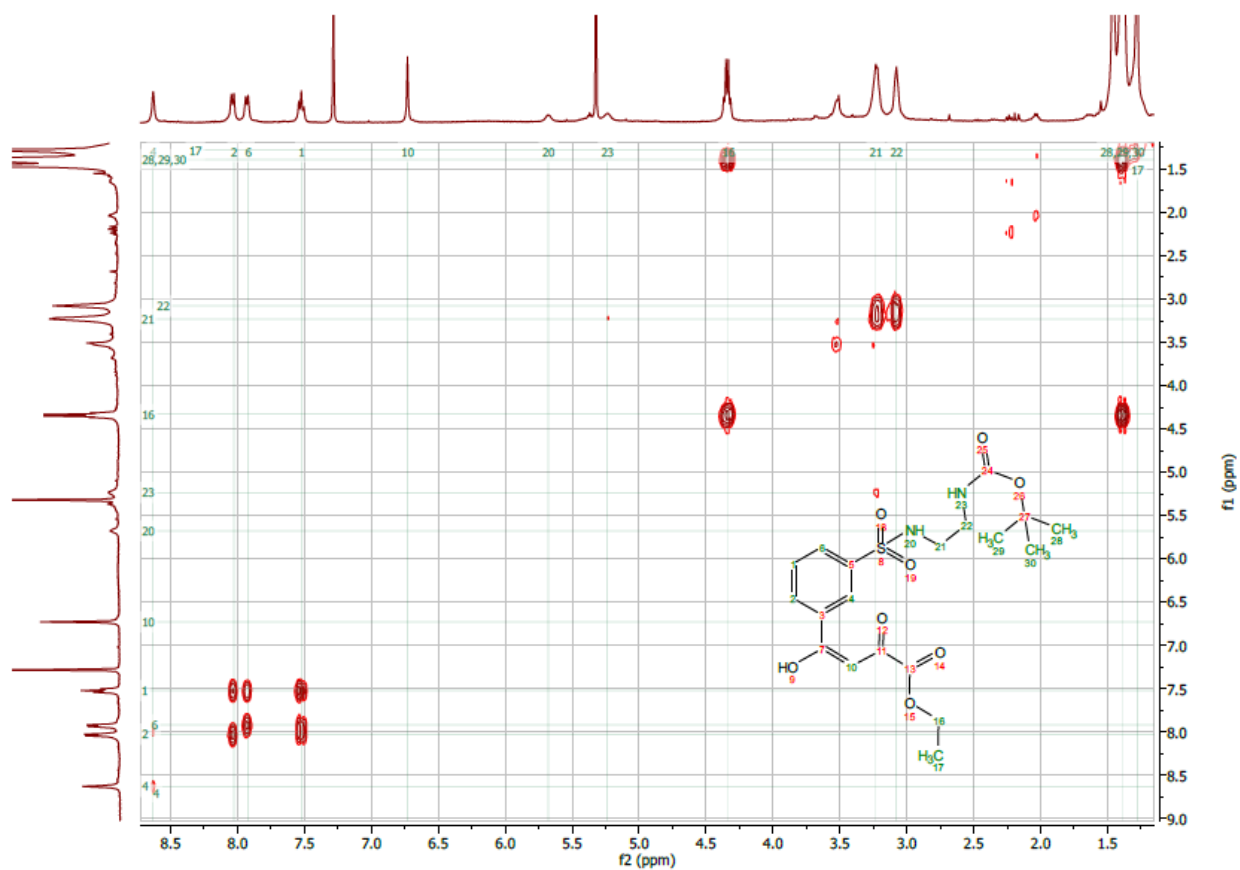
<sup>1</sup>H-NMR Ethyl 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **31b**



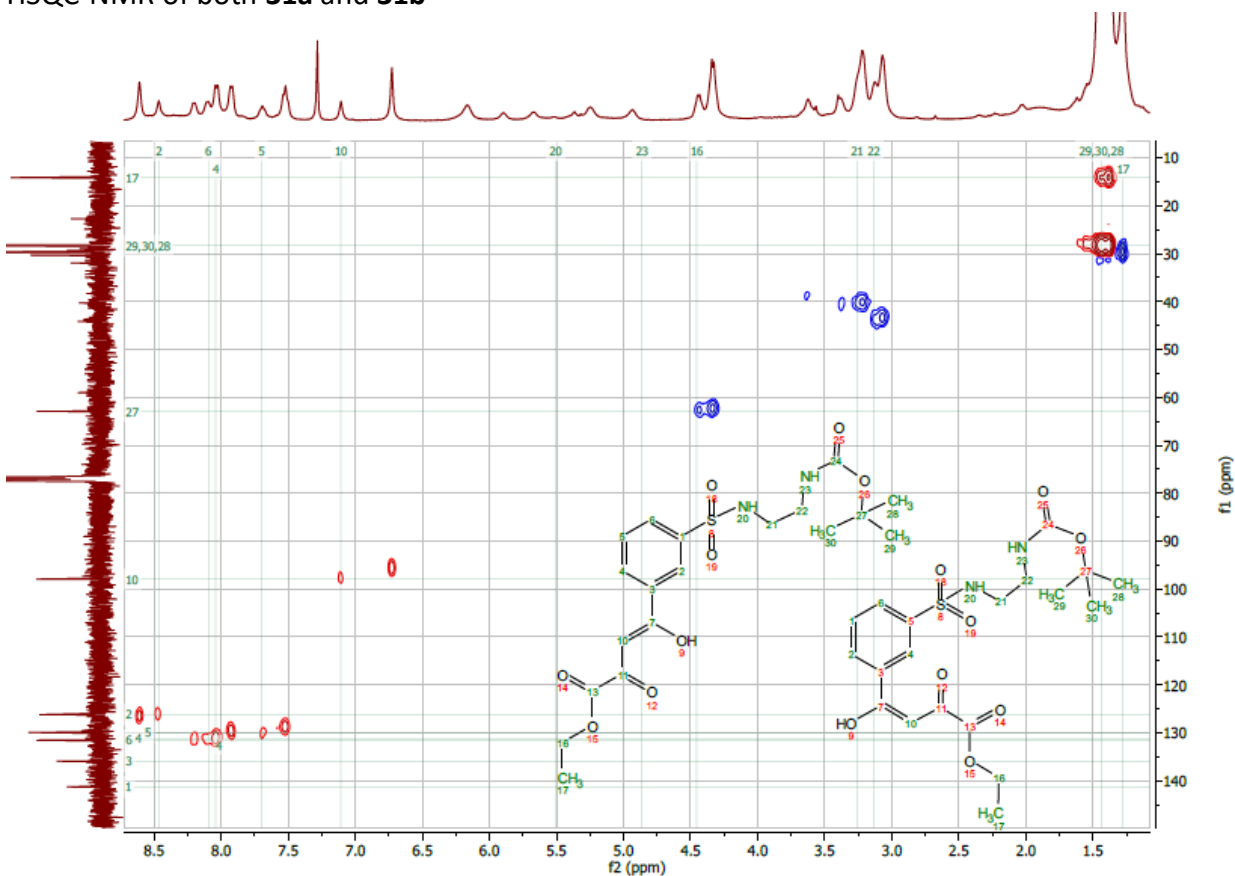
<sup>13</sup>C-NMR Ethyl 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **31b**



COSY-NMR Ethyl 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **31a**



HSQC-NMR of both **31a** and **31b**



Chemical structure of 4-aminobenzoic acid (p-aminobenzoic acid) is shown with atom numbering:

NC(=O)c1ccc(cc1)C(=O)O

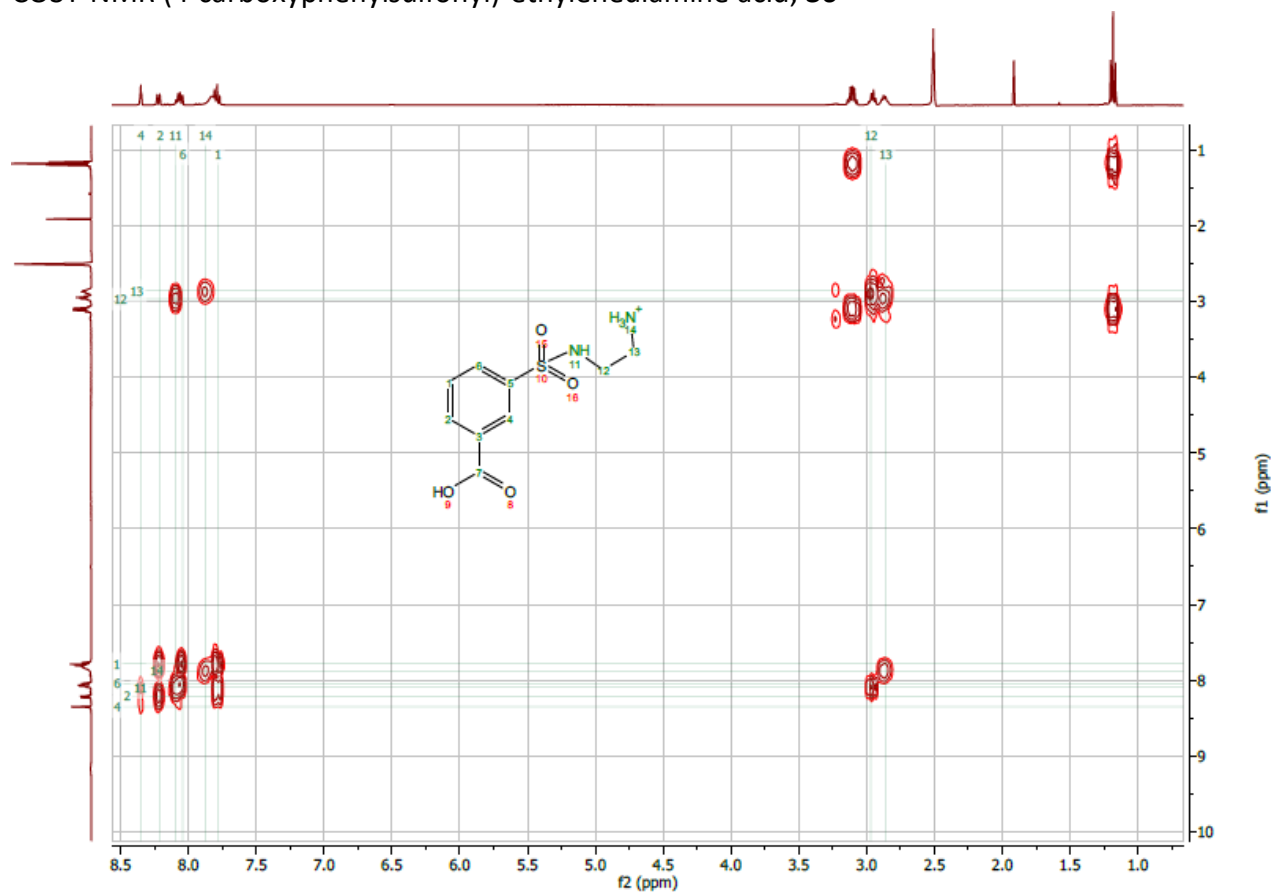
The structure is labeled with atoms 1 through 14. The carboxylic acid group is labeled 1-4, the benzene ring is labeled 5-8, and the amine group is labeled 9-14.

The  $^1\text{H}$  NMR spectrum (DMSO- $d_6$ ) shows the following peaks and integrations:

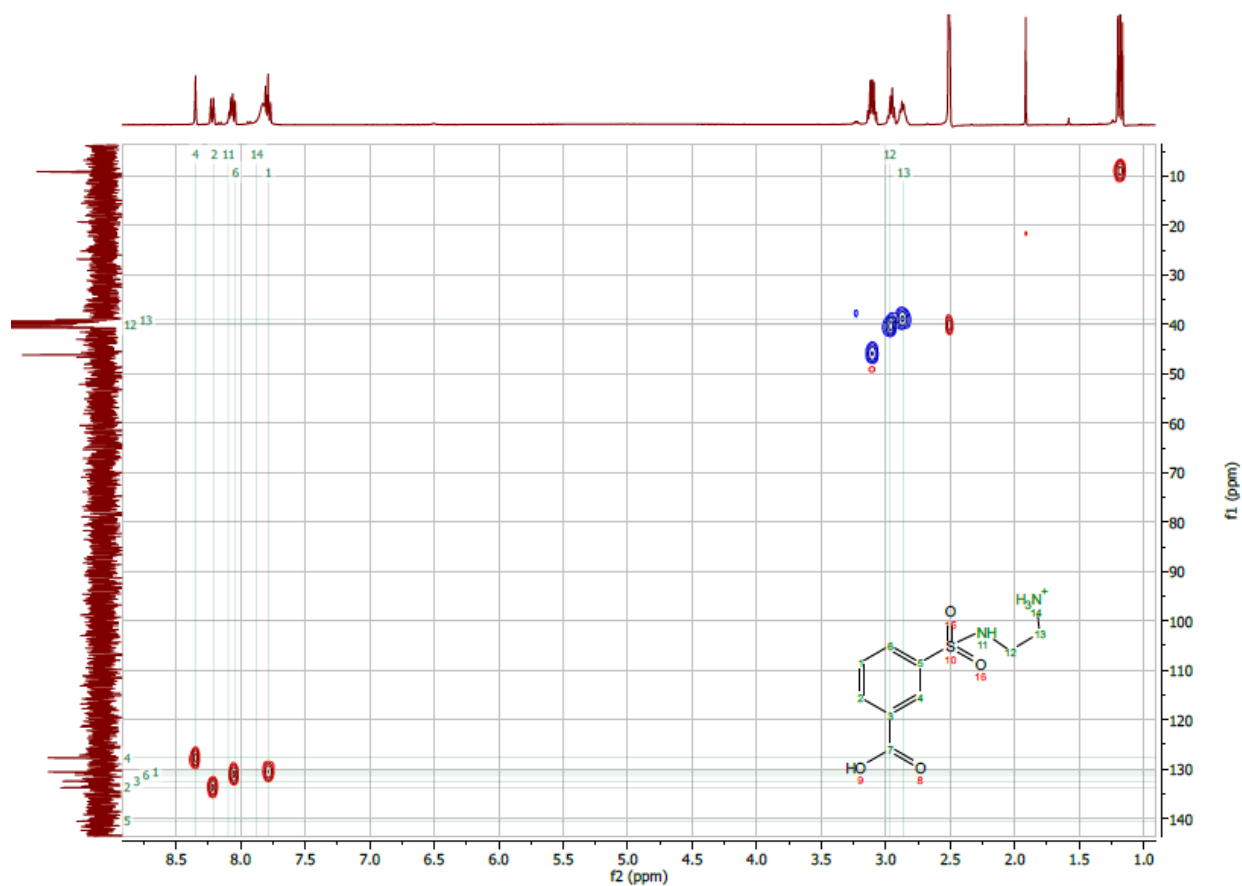
Chemical Shift (ppm)	Integration
12.0	1.00
8.35	1.00
8.23	1.00
8.11	1.00
8.09	1.00
8.07	1.00
8.04	1.00
8.04	1.00
7.76	1.00
7.76	1.00
2.52	2.27
2.51	2.36
2.50	2.36

Chemical structure of 4-aminobenzoic acid (p-aminobenzoic acid) is shown. The structure is a benzene ring with a carboxylic acid group (-COOH) at position 1 and an amino group (-NH<sub>2</sub>) at position 4. The carbon atoms are numbered 1 through 13, and the nitrogen atom is numbered 14. The spectrum shows peaks at 168.56, 160.55, 133.77, 132.45, 131.17, 130.58, 127.74, 127.74, 46.17, 40.62, 40.41, 40.30, 39.99, 39.78, 39.56, 39.05, and 9.06 ppm.

COSY-NMR (4-carboxyphenylsulfonyl)-ethylenediamine acid, **36**

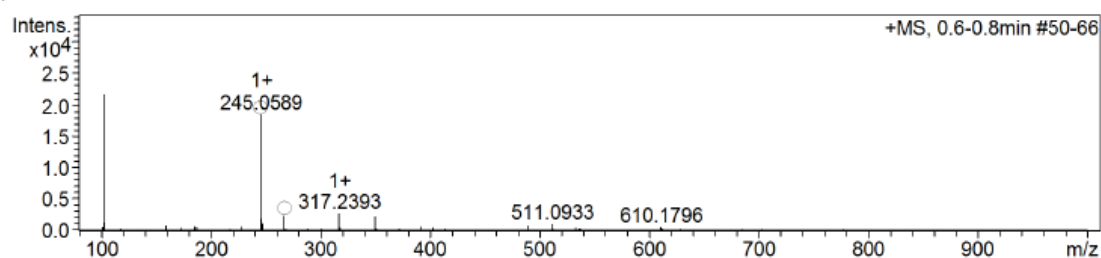


HSQC-NMR (4-carboxyphenylsulfonyl)-ethylenediamine acid **36**

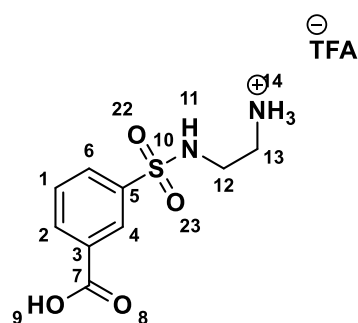


# ESI-MS (4-carboxyphenylsulfonyl)-ethylenediamine acid, **36**

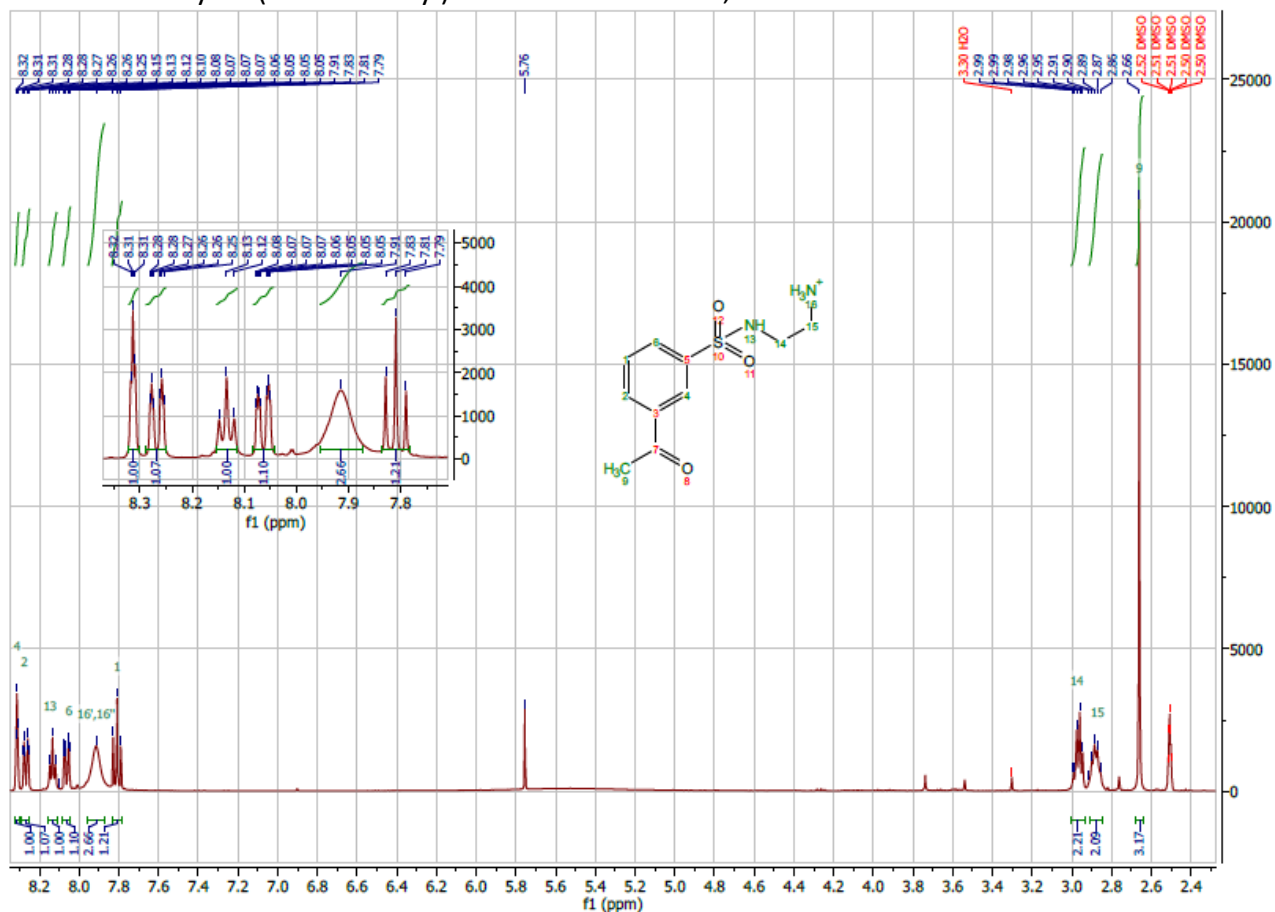
+MS, 0.6-0.8min #50-66



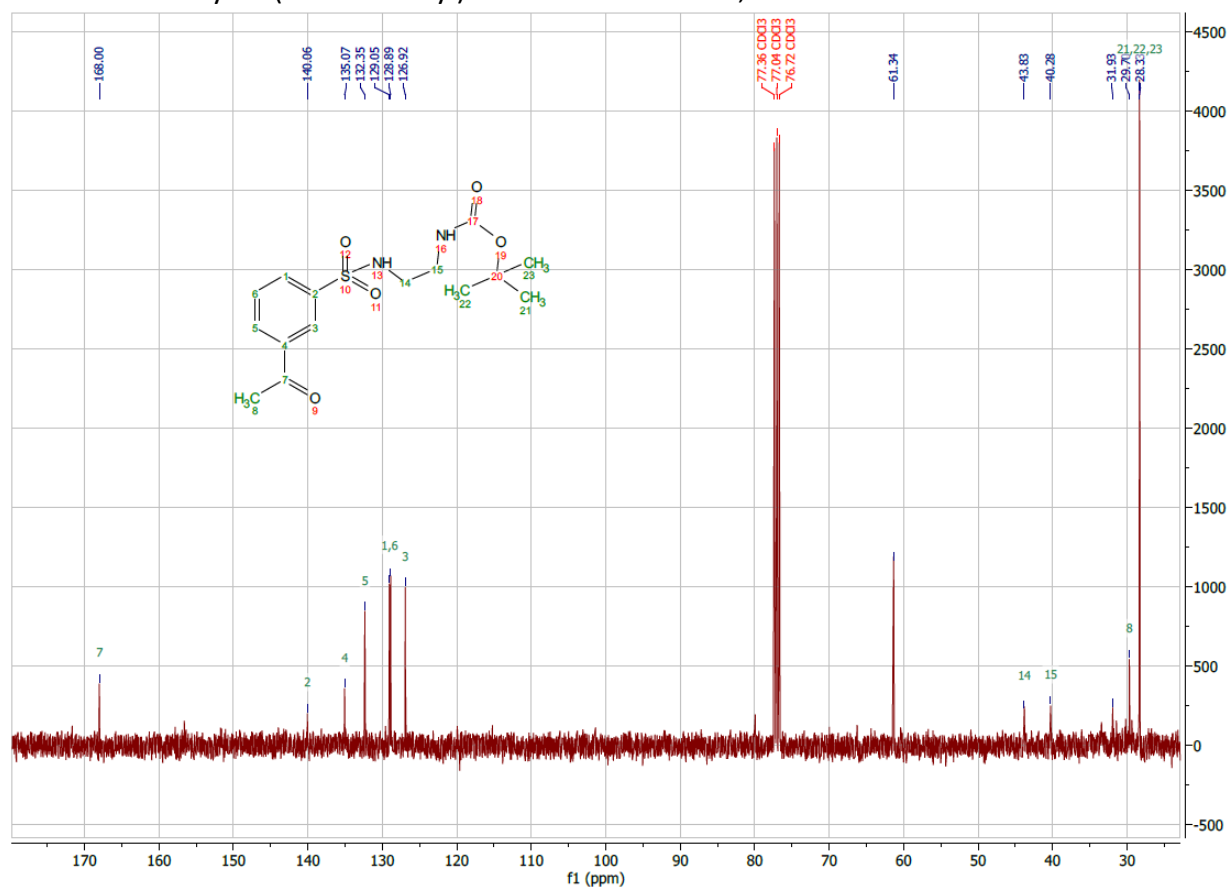
#	m/z	I %
1	101.0026	2.7
2	102.1284	100.0
3	103.1312	6.5
4	158.9638	4.1
5	173.0775	2.2
6	185.1131	3.3
7	187.0978	2.5
8	228.0321	3.3
9	245.0589	84.9
10	246.0617	9.1
11	247.0564	5.3
12	267.0395	11.4
13	317.2393	12.5
14	350.1452	10.7
15	392.1557	3.1
16	403.0170	2.4
17	489.1106	4.0
18	511.0933	5.0
19	533.0742	2.4
20	610.1796	2.7



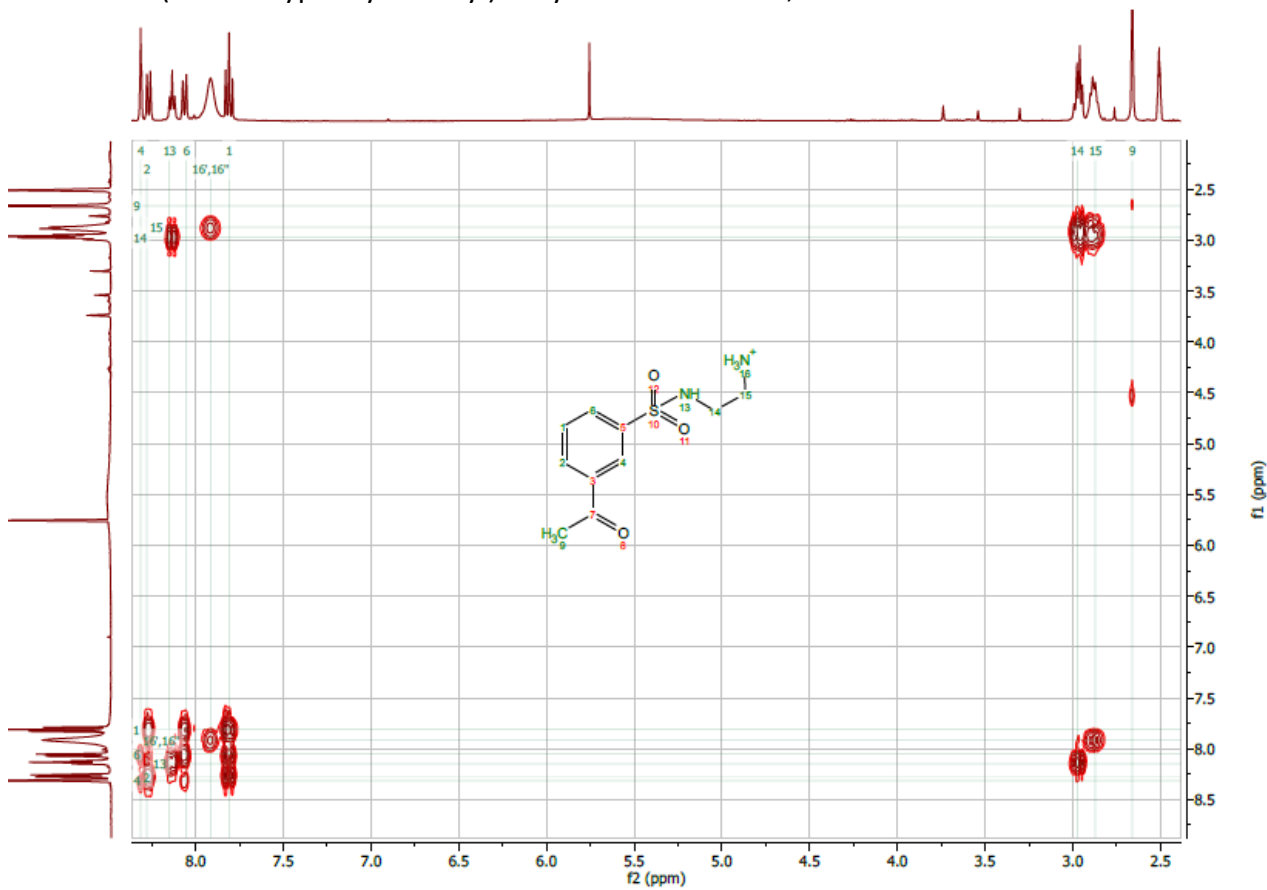
# $^1\text{H}$ -NMR 3-acetyl-N-(2-aminoethyl)benzenesulfonamide, **37**



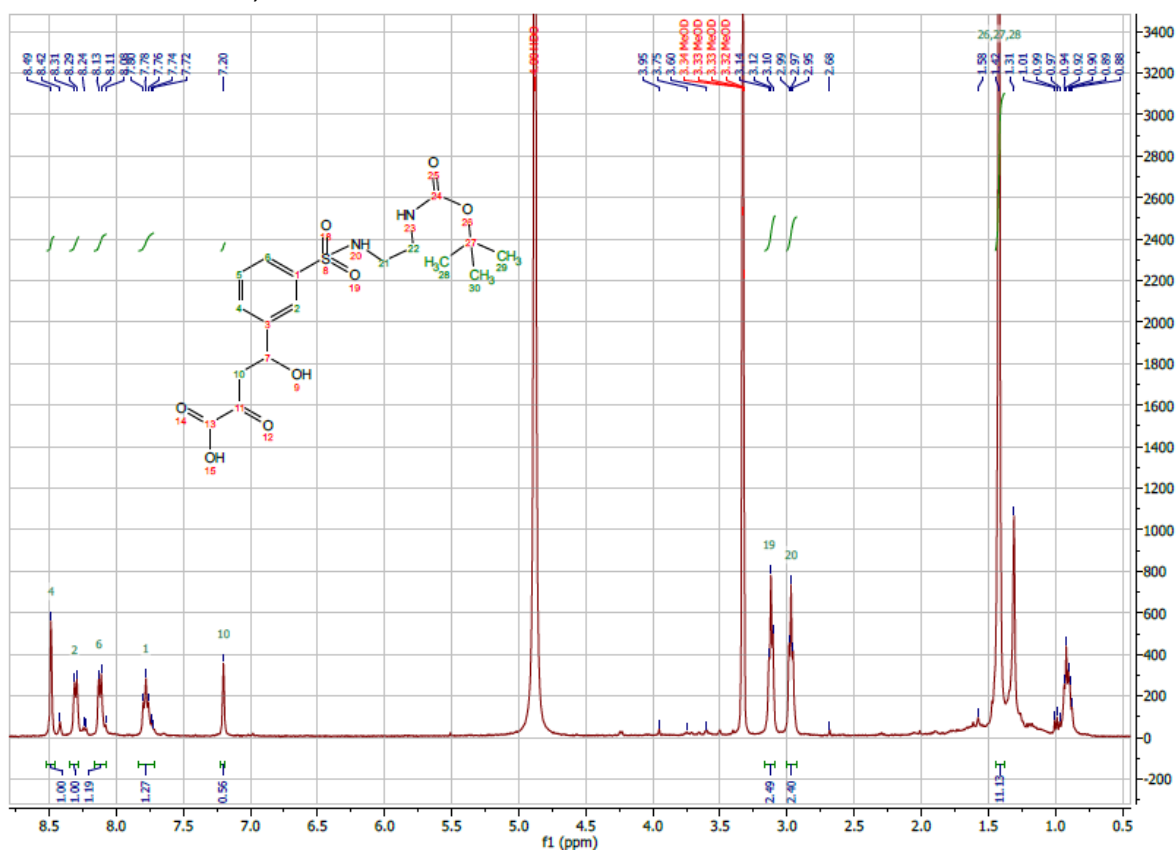
<sup>13</sup>C-NMR 3-acetyl-N-(2-aminoethyl)benzenesulfonamide, **37**



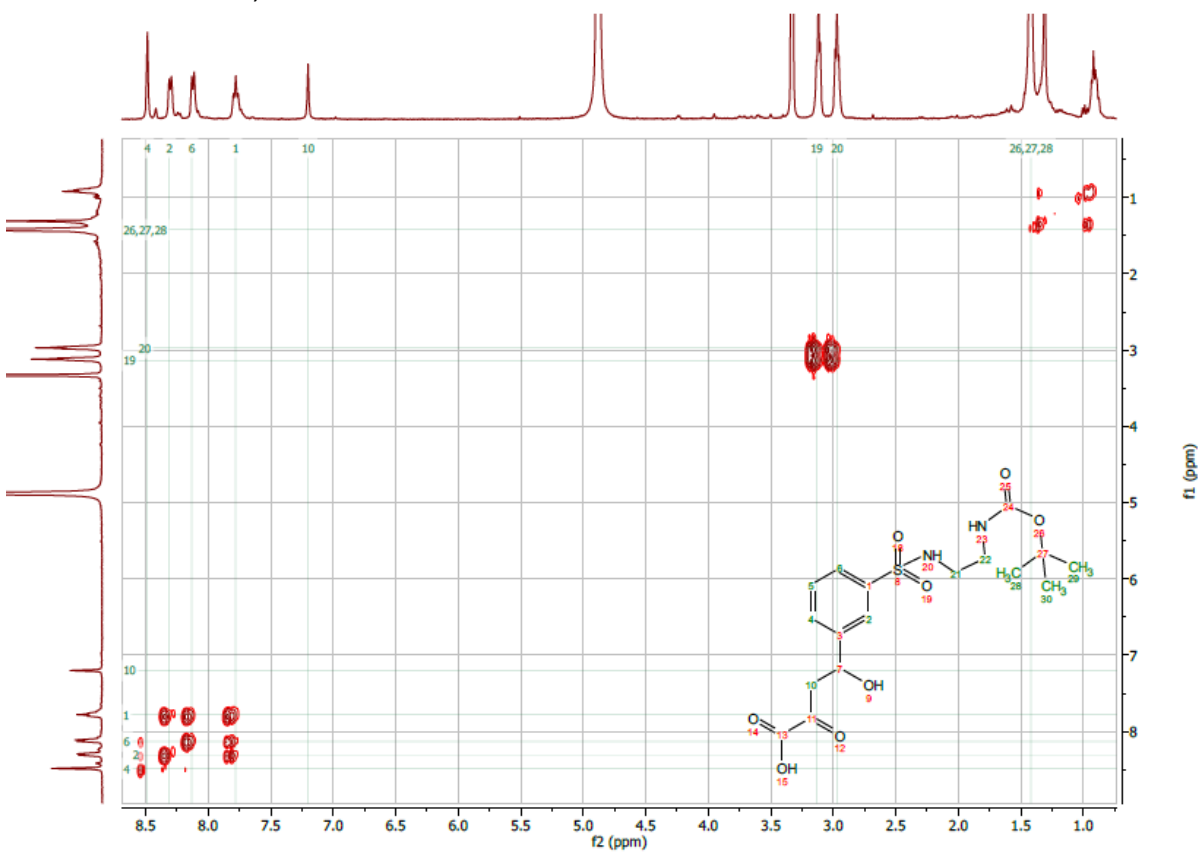
COSY-NMR (4-carboxyphenylsulfonyl)-ethylenediamine acid, **37**



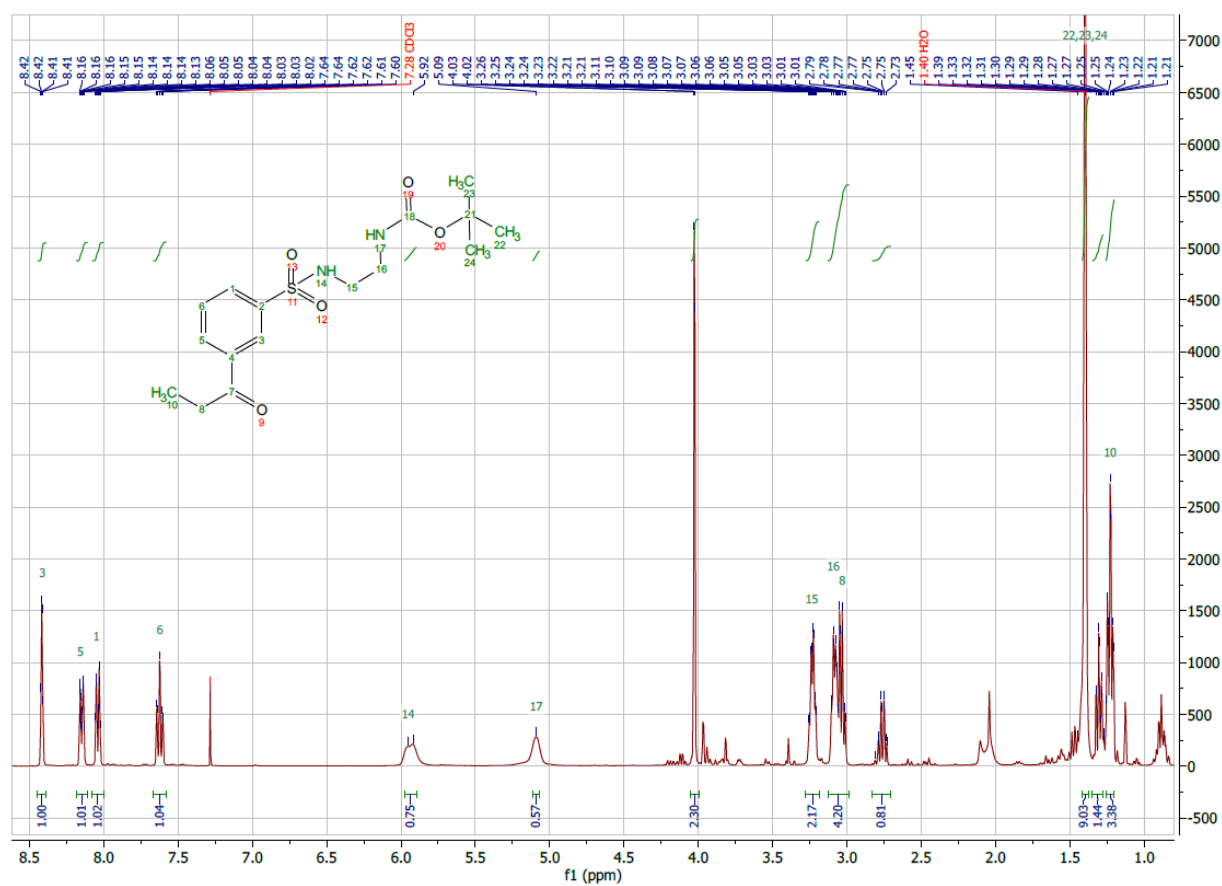
<sup>1</sup>H-NMR 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoic acid, **38**



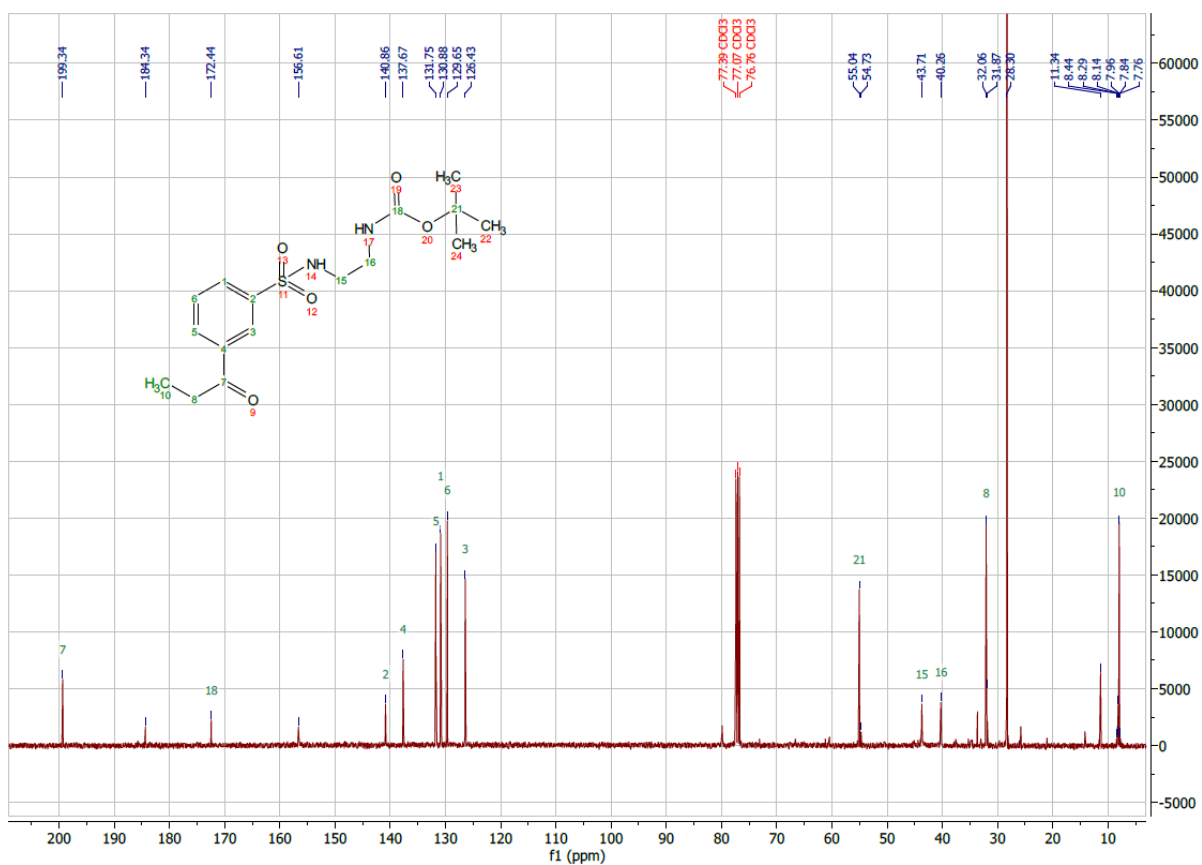
COSY-NMR 4-(3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoic acid, **38**



<sup>1</sup>H-NMR *tert*-butyl (2-((3-propionylphenyl)sulfonamido)ethyl)carbamate, **39**

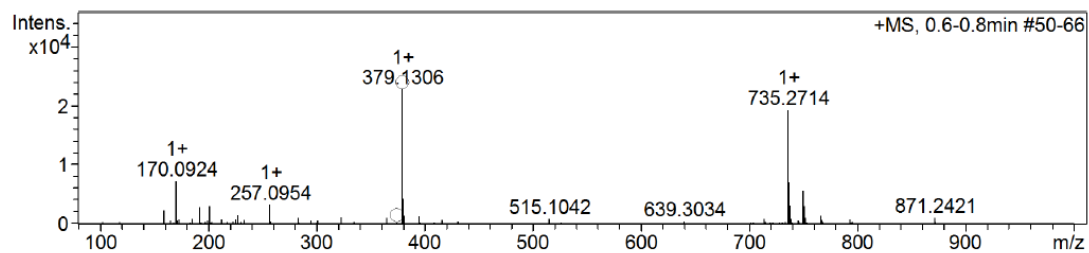


<sup>13</sup>C-NMR *tert*-butyl (2-((3-propionylphenyl)sulfonamido)ethyl)carbamate, **39**



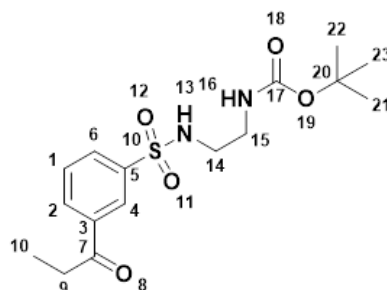


ESI-MS *tert*-butyl 2-((3-propionylphenyl)sulfonamido)ethyl)carbamate, **39**

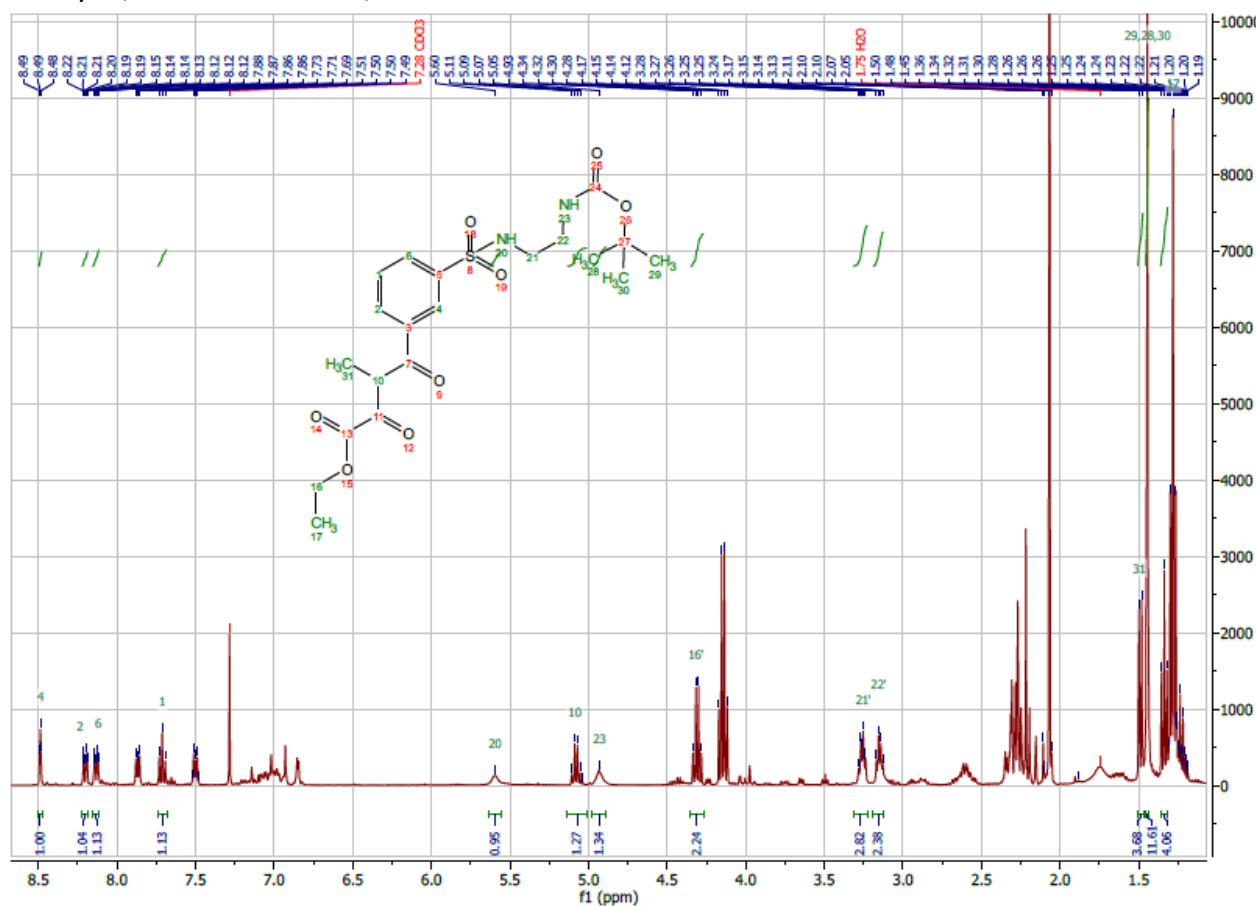


#	m/z	I %
1	158.9641	10.5
2	170.0924	32.1
3	192.0738	12.6
4	201.0976	13.5
5	227.1141	6.7
6	257.0954	14.4
7	283.1638	4.9
8	323.0668	5.5
9	365.1057	4.8
10	379.1306	100.0
11	380.1326	18.8
12	381.1292	6.3
13	395.1033	5.8
14	735.2714	84.5
15	736.2735	31.0
16	737.2719	13.8
17	749.5492	24.9
18	750.5527	13.4
19	751.5523	4.9
20	765.5231	6.5

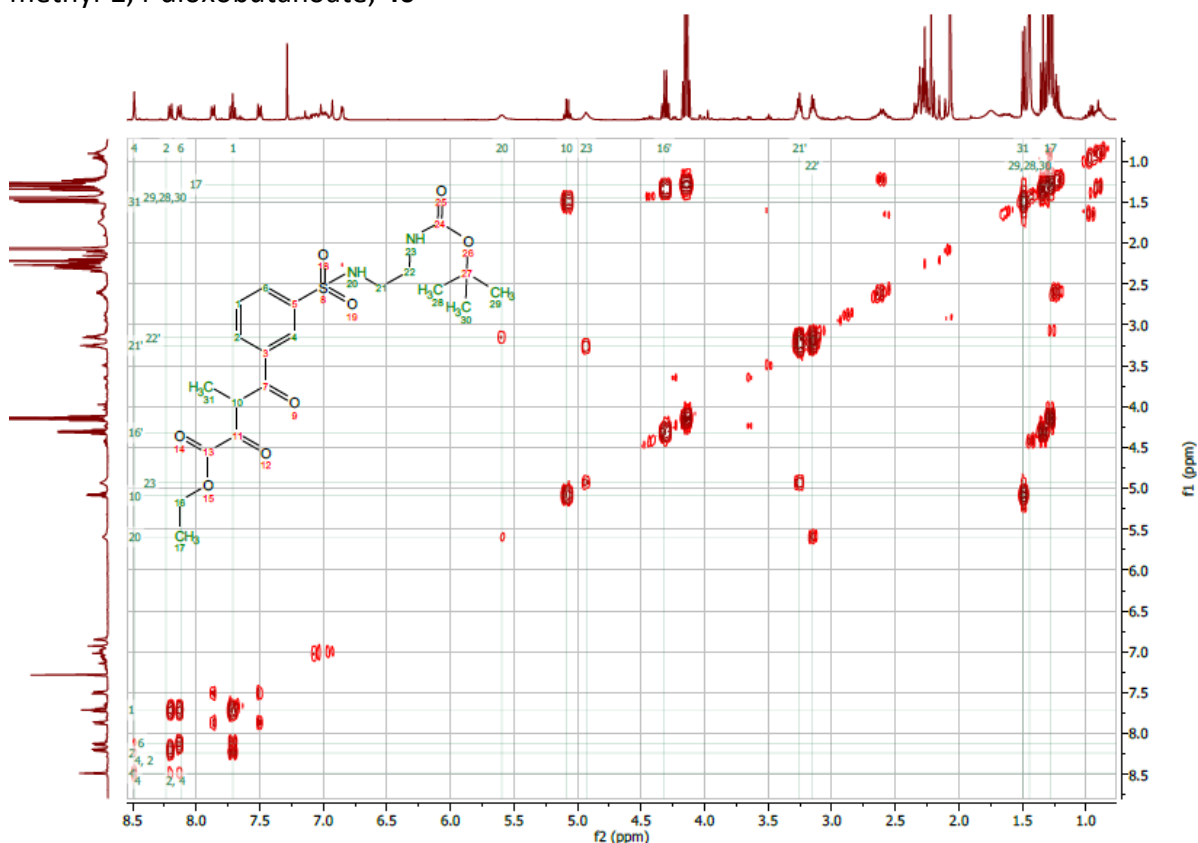
[M+Na<sup>+</sup>]



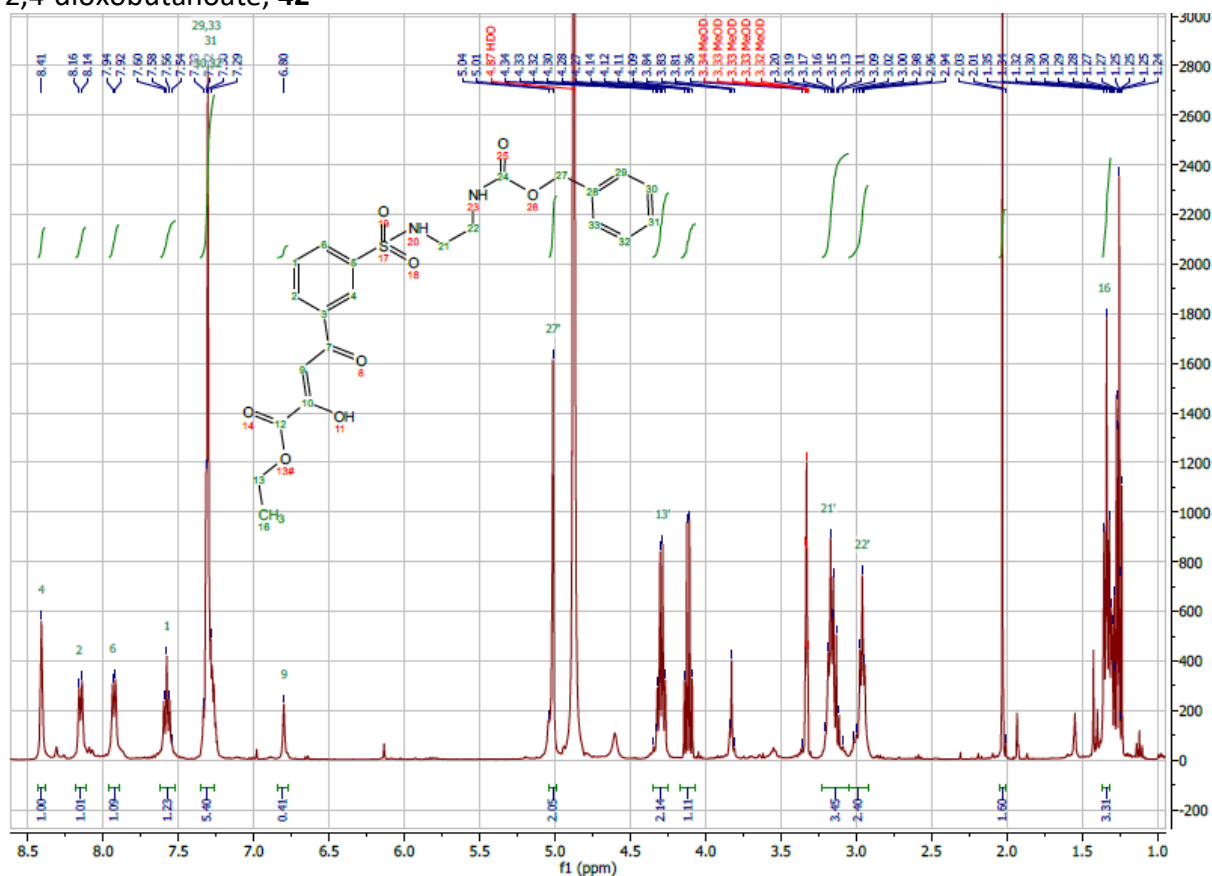
<sup>1</sup>H-NMR Ethyl 4-(3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-3-methyl-2,4-dioxobutanoate, **40**



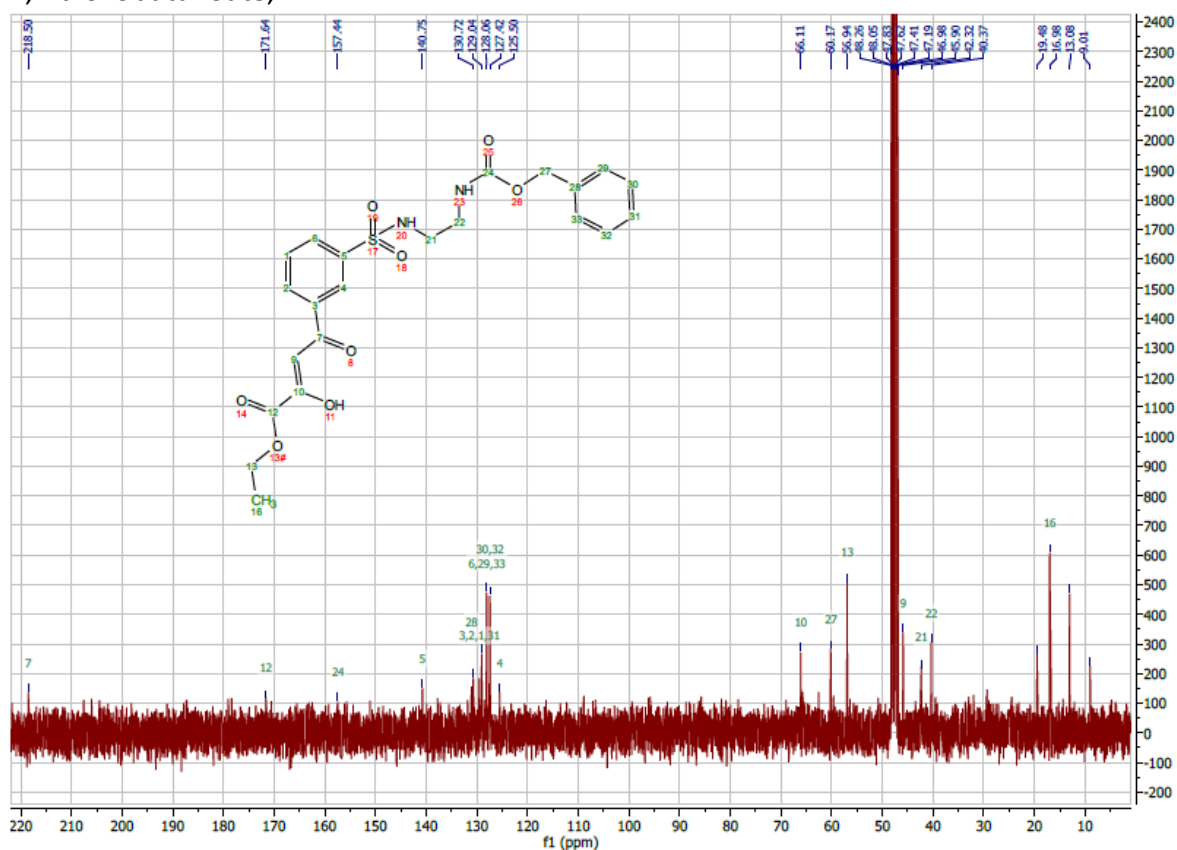
COSY-NMR Ethyl 4-(3-(N-(2-((*tert*-butoxycarbonyl)amino)ethyl)sulfamoyl)phenyl)-3-methyl-2,4-dioxobutanoate, **40**



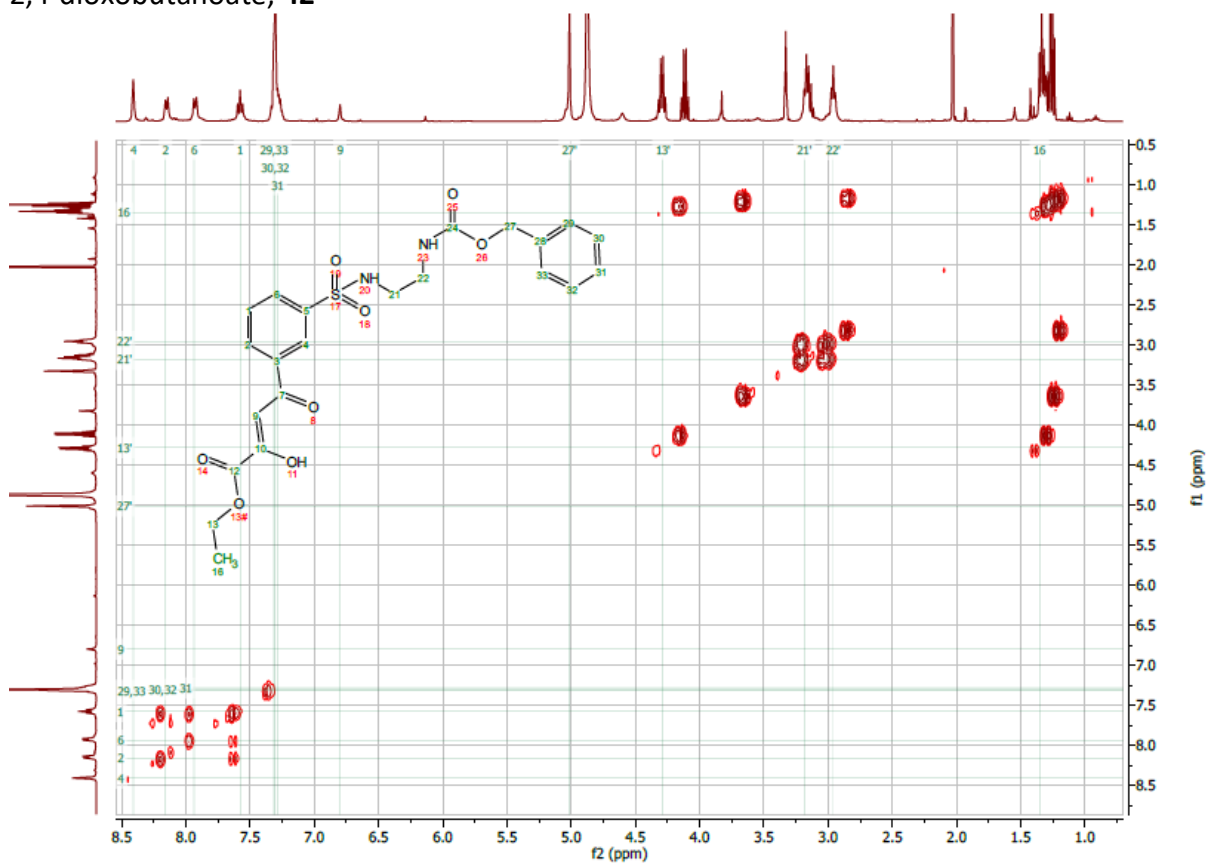
<sup>1</sup>H-NMR Methyl 4-(3-(N-(2-(((benzyloxy)carbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **42**



$^{13}\text{C}$ -NMR Methyl 4-(3-(N-(2-(((benzyloxy)carbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **42**



COSY-NMR Methyl 4-(3-(N-(2-(((benzyloxy)carbonyl)amino)ethyl)sulfamoyl)phenyl)-2,4-dioxobutanoate, **42**

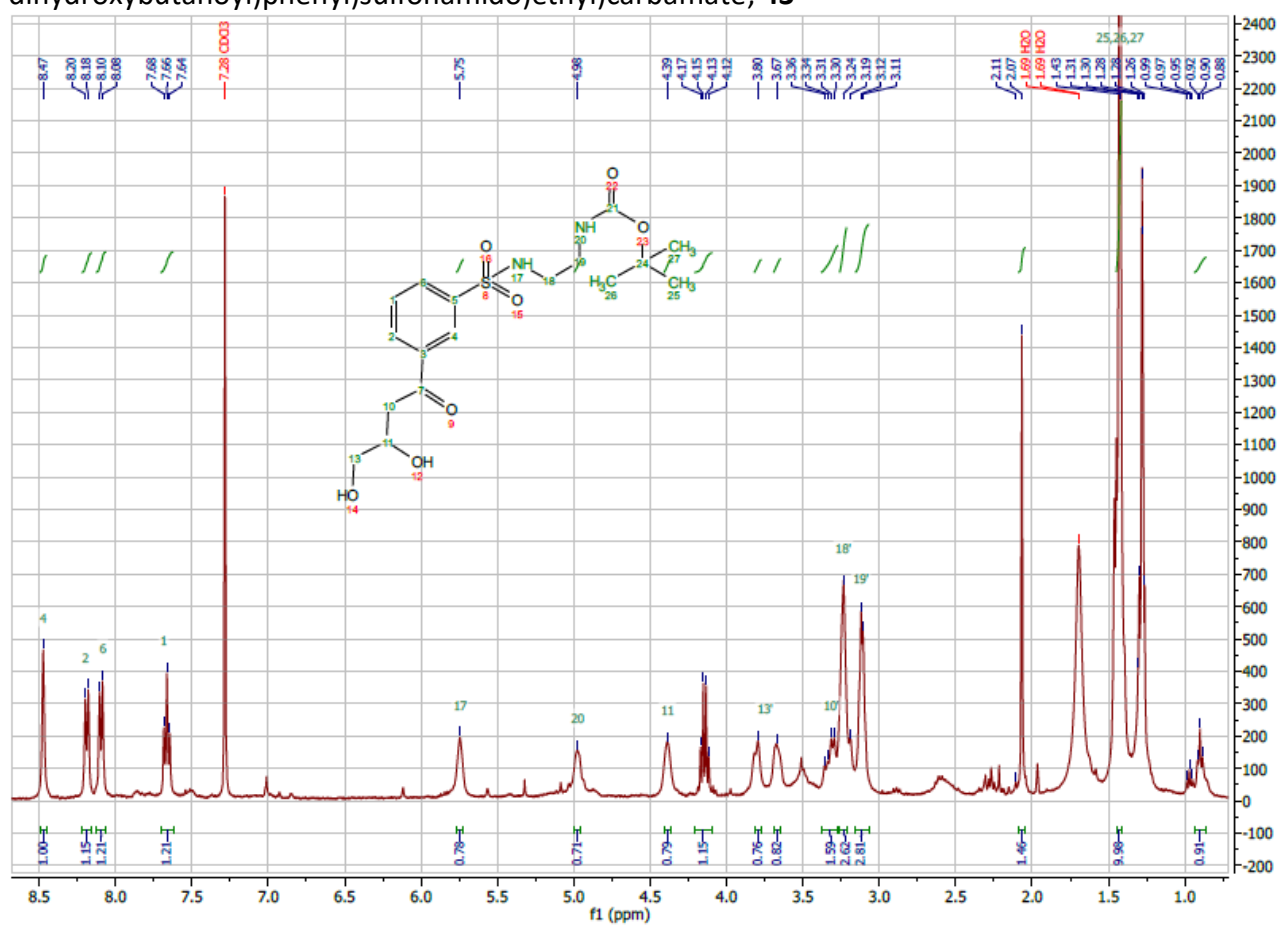


<sup>1</sup>H-NMR

*tert*-butyl

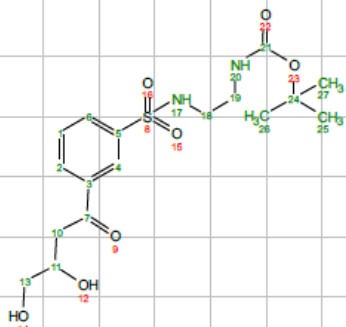
(2-((3-(3,4-

dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **43**



<sup>13</sup>C-NMR *tert*-butyl  
dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **43**

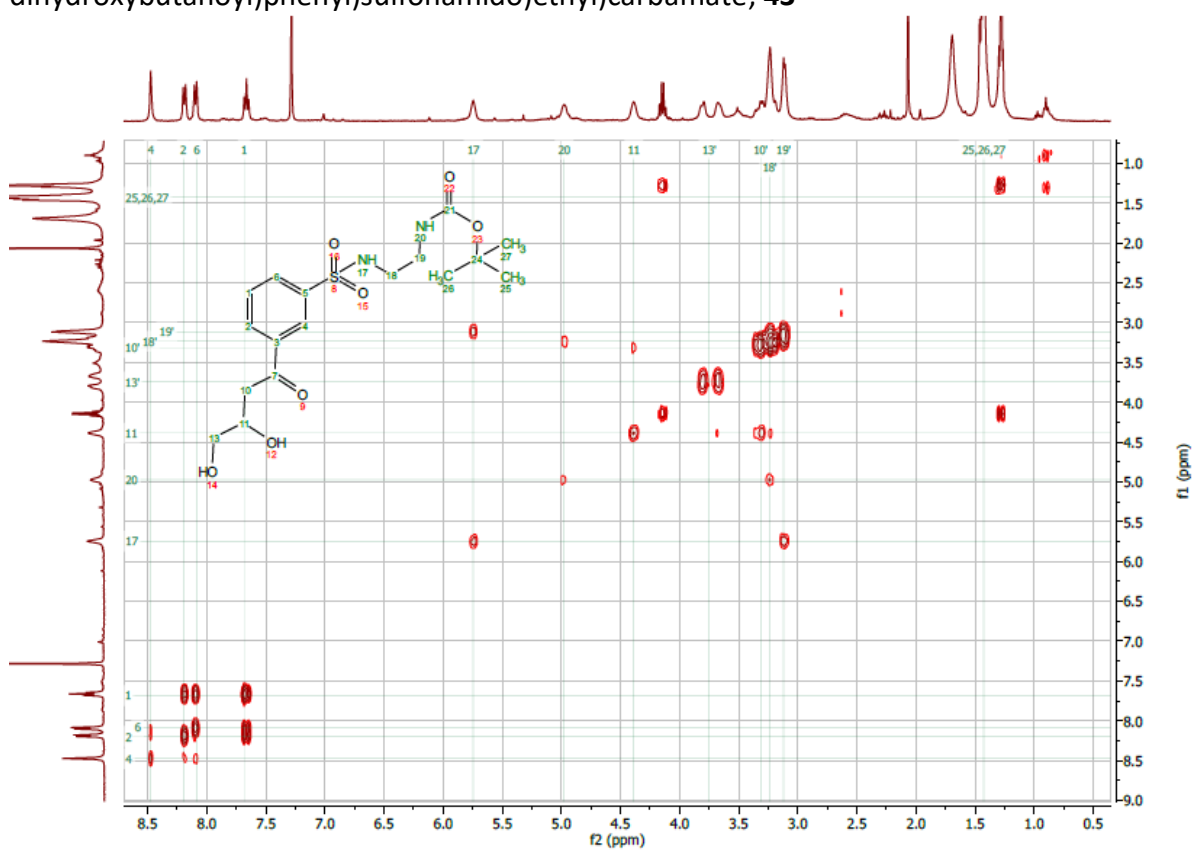
(2-((3-(3,4-



COSY-NMR *tert*-butyl  
dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **43**

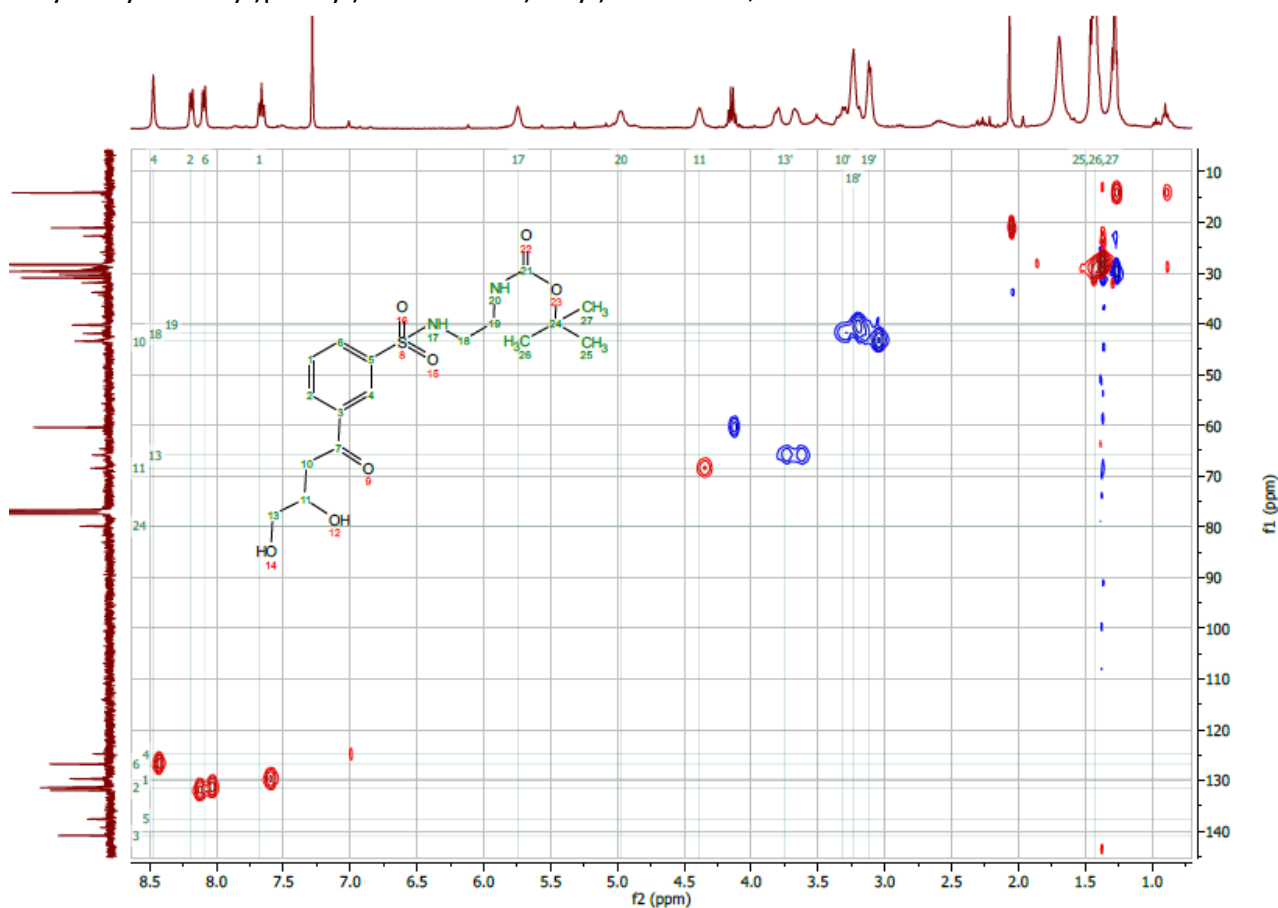
*tert*-butyl

(2-((3-(3,4-

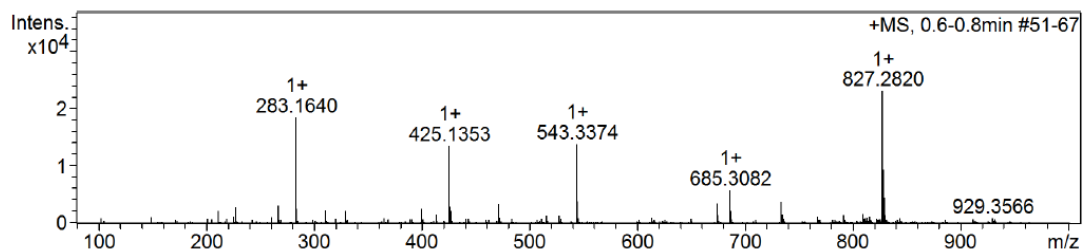


HSQC-NMR *tert*-butyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **43**

(2-((3-(3,4-



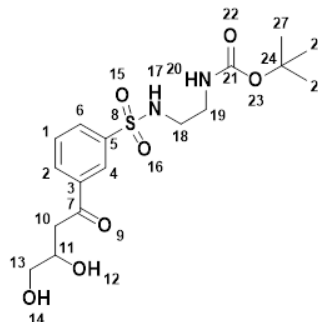
*tert*-butyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **43**



#	m/z	I %
1	227.1118	11.8
2	267.0802	13.5
3	283.1640	79.9
4	284.1657	10.7
5	311.0701	9.5
6	329.2184	9.1
7	400.2262	11.1
8	425.1353	58.5
9	426.1395	12.5
10	427.2096	9.1
11	471.2752	14.4
12	543.3374	59.8
13	544.3393	17.0
14	673.4401	15.2
15	685.3082	25.3
16	686.3114	9.3
17	733.2570	16.5
18	827.2820	100.0
19	828.2843	40.4
20	829.2946	19.4

← [M+Na<sup>+</sup>]

← [2M+Na<sup>+</sup>]



**<sup>1</sup>H NMR Spectrum (DMSO-d<sub>6</sub>) of 4-aminobenzoic acid**

**Chemical Structure:** NC(=O)c1ccc(cc1)C(=O)O

**Peak Data:**

Chemical Shift (ppm)	Integration
7.819, 7.818, 7.817, 7.800, 7.799, 7.798, 7.797, 7.779, 7.778, 7.777, 7.766, 7.666, 7.665, 7.664, 7.662, 7.661, 7.660, 7.659, 7.658, 7.657, 7.656, 7.655, 7.654, 7.653, 7.652, 7.651, 7.650, 7.649, 7.648, 7.647, 7.646, 7.645, 7.644, 7.643, 7.642, 7.641, 7.640, 7.639, 7.638, 7.637, 7.636, 7.635, 7.634, 7.633, 7.632, 7.631, 7.630, 7.629, 7.628, 7.627, 7.626, 7.625, 7.624, 7.623, 7.622, 7.621, 7.620, 7.619, 7.618, 7.617, 7.616, 7.615, 7.614, 7.613, 7.612, 7.611, 7.610, 7.609, 7.608, 7.607, 7.606, 7.605, 7.604, 7.603, 7.602, 7.601, 7.600, 7.599, 7.598, 7.597, 7.596, 7.595, 7.594, 7.593, 7.592, 7.591, 7.590, 7.589, 7.588, 7.587, 7.586, 7.585, 7.584, 7.583, 7.582, 7.581, 7.580, 7.579, 7.578, 7.577, 7.576, 7.575, 7.574, 7.573, 7.572, 7.571, 7.570, 7.569, 7.568, 7.567, 7.566, 7.565, 7.564, 7.563, 7.562, 7.561, 7.560, 7.559, 7.558, 7.557, 7.556, 7.555, 7.554, 7.553, 7.552, 7.551, 7.550, 7.549, 7.548, 7.547, 7.546, 7.545, 7.544, 7.543, 7.542, 7.541, 7.540, 7.539, 7.538, 7.537, 7.536, 7.535, 7.534, 7.533, 7.532, 7.531, 7.530, 7.529, 7.528, 7.527, 7.526, 7.525, 7.524, 7.523, 7.522, 7.521, 7.520, 7.519, 7.518, 7.517, 7.516, 7.515, 7.514, 7.513, 7.512, 7.511, 7.510, 7.509, 7.508, 7.507, 7.506, 7.505, 7.504, 7.503, 7.502, 7.501, 7.500, 7.499, 7.498, 7.497, 7.496, 7.495, 7.494, 7.493, 7.492, 7.491, 7.490, 7.489, 7.488, 7.487, 7.486, 7.485, 7.484, 7.483, 7.482, 7.481, 7.480, 7.479, 7.478, 7.477, 7.476, 7.475, 7.474, 7.473, 7.472, 7.471, 7.470, 7.469, 7.468, 7.467, 7.466, 7.465, 7.464, 7.463, 7.462, 7.461, 7.460, 7.459, 7.458, 7.457, 7.456, 7.455, 7.454, 7.453, 7.452, 7.451, 7.450, 7.449, 7.448, 7.447, 7.446, 7.445, 7.444, 7.443, 7.442, 7.441, 7.440, 7.439, 7.438, 7.437, 7.436, 7.435, 7.434, 7.433, 7.432, 7.431, 7.430, 7.429, 7.428, 7.427, 7.426, 7.425, 7.424, 7.423, 7.422, 7.421, 7.420, 7.419, 7.418, 7.417, 7.416, 7.415, 7.414, 7.413, 7.412, 7.411, 7.410, 7.409, 7.408, 7.407, 7.406, 7.405, 7.404, 7.403, 7.402, 7.401, 7.400, 7.399, 7.398, 7.397, 7.396, 7.395, 7.394, 7.393, 7.392, 7.391, 7.390, 7.389, 7.388, 7.387, 7.386, 7.385, 7.384, 7.383, 7.382, 7.381, 7.380, 7.379, 7.378, 7.377, 7.376, 7.375, 7.374, 7.373, 7.372, 7.371, 7.370, 7.369, 7.368, 7.367, 7.366, 7.365, 7.364, 7.363, 7.362, 7.361, 7.360, 7.359, 7.358, 7.357, 7.356, 7.355, 7.354, 7.353, 7.352, 7.351, 7.350, 7.349, 7.348, 7.347, 7.346, 7.345, 7.344, 7.343, 7.342, 7.341, 7.340, 7.339, 7.338, 7.337, 7.336, 7.335, 7.334, 7.333, 7.332, 7.331, 7.330, 7.329, 7.328, 7.327, 7.326, 7.325, 7.324, 7.323, 7.322, 7.321, 7.320, 7.319, 7.318, 7.317, 7.316, 7.315, 7.314, 7.313, 7.312, 7.311, 7.310, 7.309, 7.308, 7.307, 7.306, 7.305, 7.304, 7.303, 7.302, 7.301, 7.300, 7.299, 7.298, 7.297, 7.296, 7.295, 7.294, 7.293, 7.292, 7.291, 7.290, 7.289, 7.288, 7.287, 7.286, 7.285, 7.284, 7.283, 7.282, 7.281, 7.280, 7.279, 7.278, 7.277, 7.276, 7.275, 7.274, 7.273, 7.272, 7.271, 7.270, 7.269, 7.268, 7.267, 7.266, 7.265, 7.264, 7.263, 7.262, 7.261, 7.260, 7.259, 7.258, 7.257, 7.256, 7.255, 7.254, 7.253, 7.252, 7.251, 7.250, 7.249, 7.248, 7.247, 7.246, 7.245, 7.244, 7.243, 7.242, 7.241, 7.240, 7.239, 7.238, 7.237, 7.236, 7.235, 7.234, 7.233, 7.232, 7.231, 7.230, 7.229, 7.228, 7.227, 7.226, 7.225, 7.224, 7.223, 7.222, 7.221, 7.220, 7.219, 7.218, 7.217, 7.216, 7.215, 7.214, 7.213, 7.212, 7.211, 7.210, 7.209, 7.208, 7.207, 7.206, 7.205, 7.204, 7.203, 7.202, 7.201, 7.200, 7.199, 7.198, 7.197, 7.196, 7.195, 7.194, 7.193, 7.192, 7.191, 7.190, 7.189, 7.188, 7.187, 7.186, 7.185, 7.184, 7.183, 7.182, 7.181, 7.180, 7.179, 7.178, 7.177, 7.176, 7.175, 7.174, 7.173, 7.172, 7.171, 7.170, 7.169, 7.168, 7.167, 7.166, 7.165, 7.164, 7.163, 7.162, 7.161, 7.160, 7.159, 7.158, 7.157, 7.156, 7.155, 7.154, 7.153, 7.152, 7.151, 7.150, 7.149, 7.148, 7.147, 7.146, 7.145, 7.144, 7.143, 7.142, 7.141, 7.140, 7.139, 7.138, 7.137, 7.136, 7.135, 7.134, 7.133, 7.132, 7.131, 7.130, 7.129, 7.128, 7.127, 7.126, 7.125, 7.124, 7.123,	

Chemical structure of 2-(furan-2-yl)-N-(2-aminophenyl)benzamide is shown. The structure is labeled with carbon numbers 1 through 17, corresponding to the peaks in the  $^{13}\text{C}$  NMR spectrum.

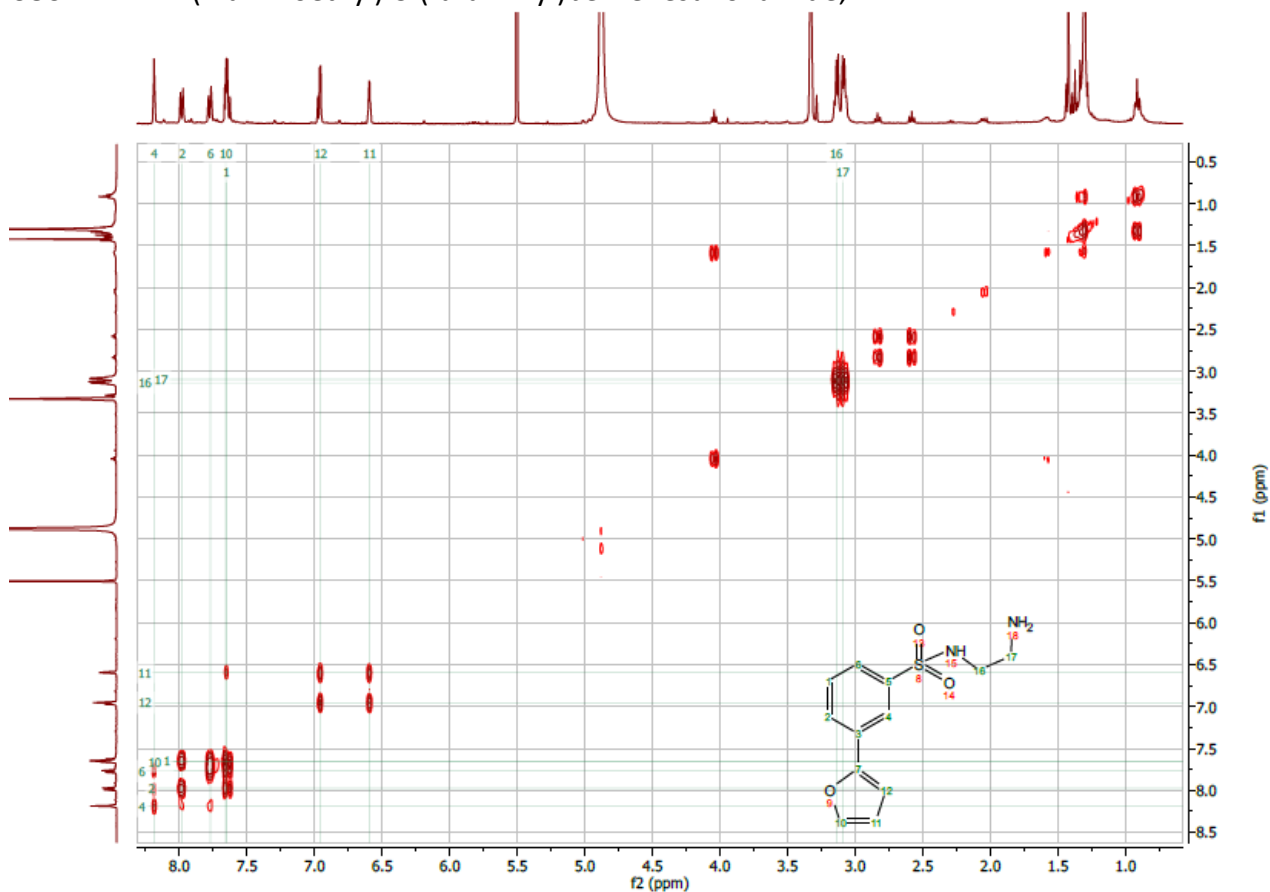
The  $^{13}\text{C}$  NMR spectrum (f1 (ppm)) displays the following chemical shifts (ppm):

- 151.94
- 145.21
- 140.29
- 132.10
- 129.96
- 127.95
- 125.07
- 121.40
- 111.76
- 106.74
- 48.24
- 48.03
- 47.82
- 47.59
- 47.18
- 46.96
- 40.00
- 39.28
- 29.35

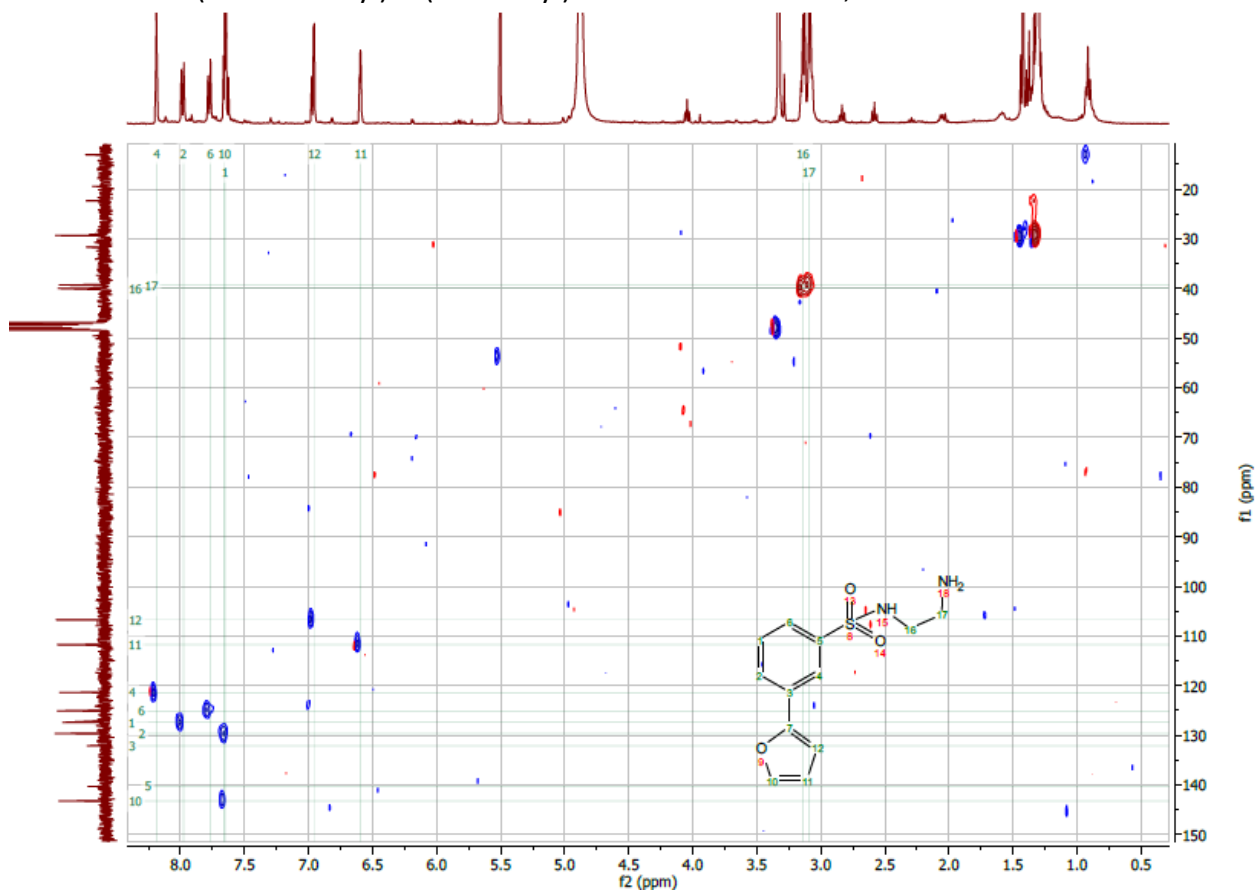
Peak assignments (Carbon Number) are indicated above the peaks:

- 7 (151.94)
- 10 (145.21)
- 5 (140.29)
- 3 (132.10)
- 2 (129.96)
- 1 (127.95)
- 6 (125.07)
- 4 (121.40)
- 11 (111.76)
- 12 (106.74)
- 17 (48.24)
- 16 (47.82)
- 15 (29.35)

COSY-NMR N-(2-aminoethyl)-3-(furan-2-yl)benzenesulfonamide, **44**



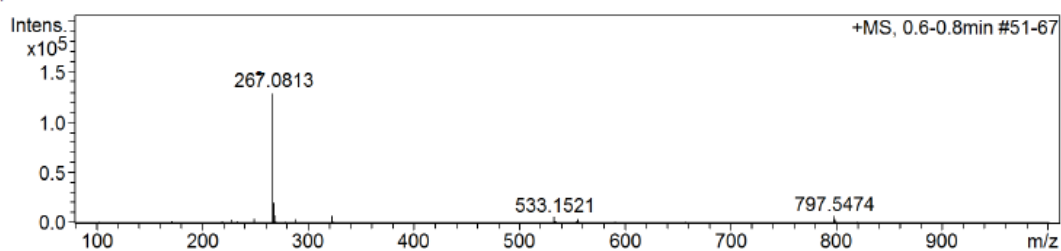
HSQC-NMR N-(2-aminoethyl)-3-(furan-2-yl)benzenesulfonamide, **44**





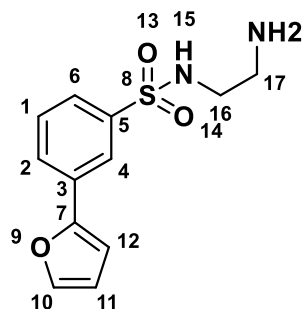
# ESI-MS N-(2-aminoethyl)-3-(furan-2-yl)benzenesulfonamide, **44**

+MS, 0.6-0.8min #51-67

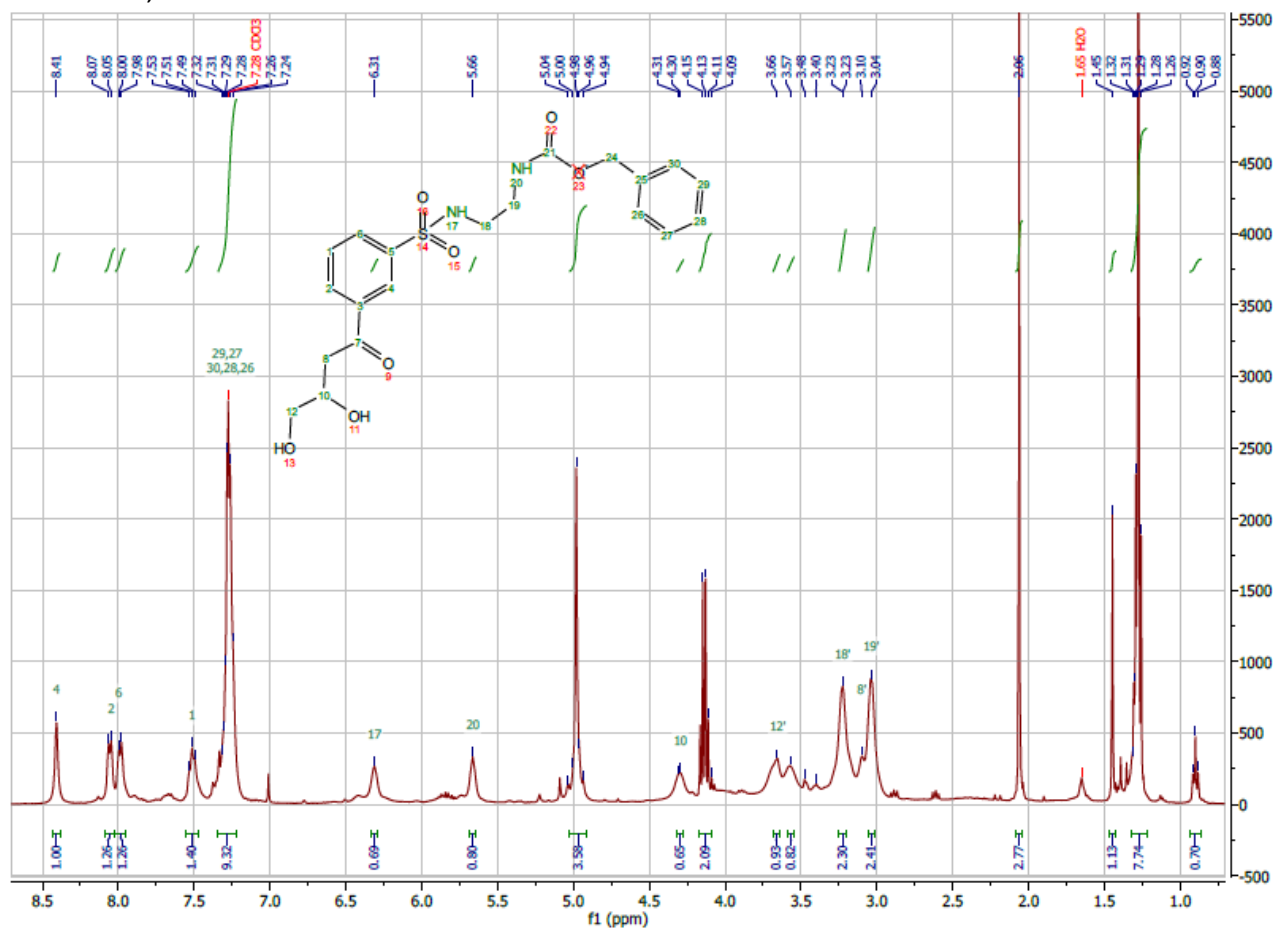


#	m/z	I %
1	171.0558	1.6
2	219.1873	1.3
3	227.1152	2.4
4	233.2030	1.4
5	250.0535	3.1
6	267.0813	100.0
7	268.0837	15.8
8	269.0787	6.0
9	270.0962	0.9
10	279.0825	1.0
11	289.0653	2.8
12	323.1433	5.4
13	324.1447	1.1
14	533.1521	4.5
15	534.1549	1.3
16	553.4605	1.2
17	555.1340	2.6
18	797.5474	5.6
19	798.5502	3.0
20	799.5531	1.1

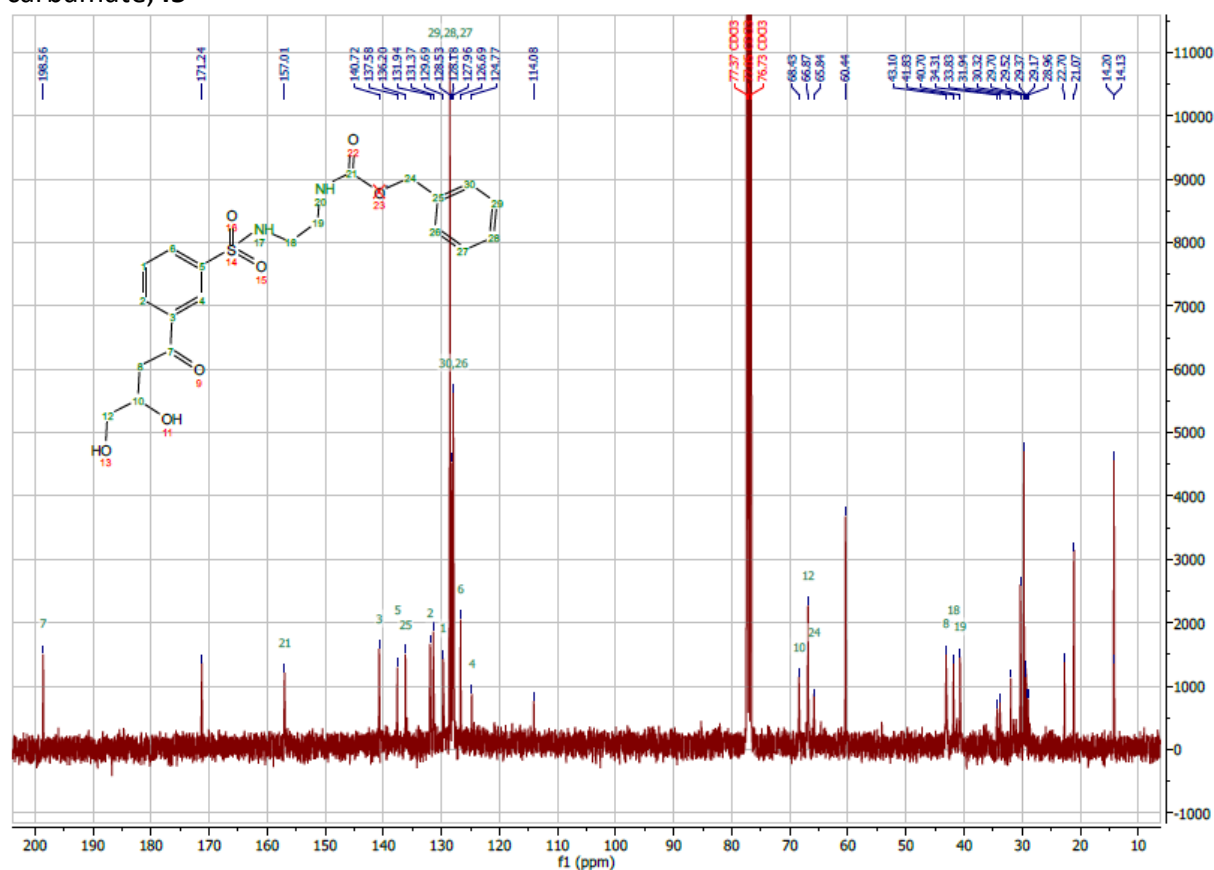
[M+H<sup>+</sup>]



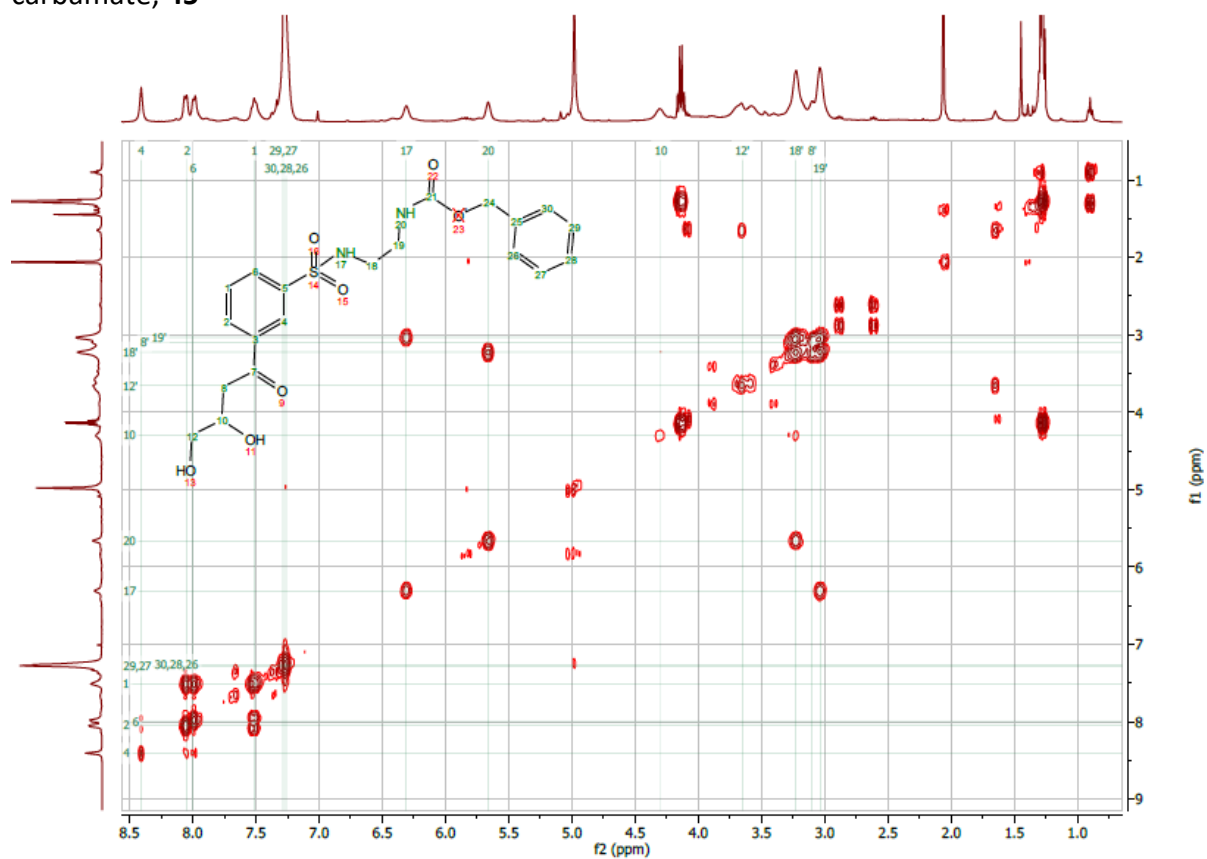
## <sup>1</sup>H-NMR Benzyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl) carbamate, **45**



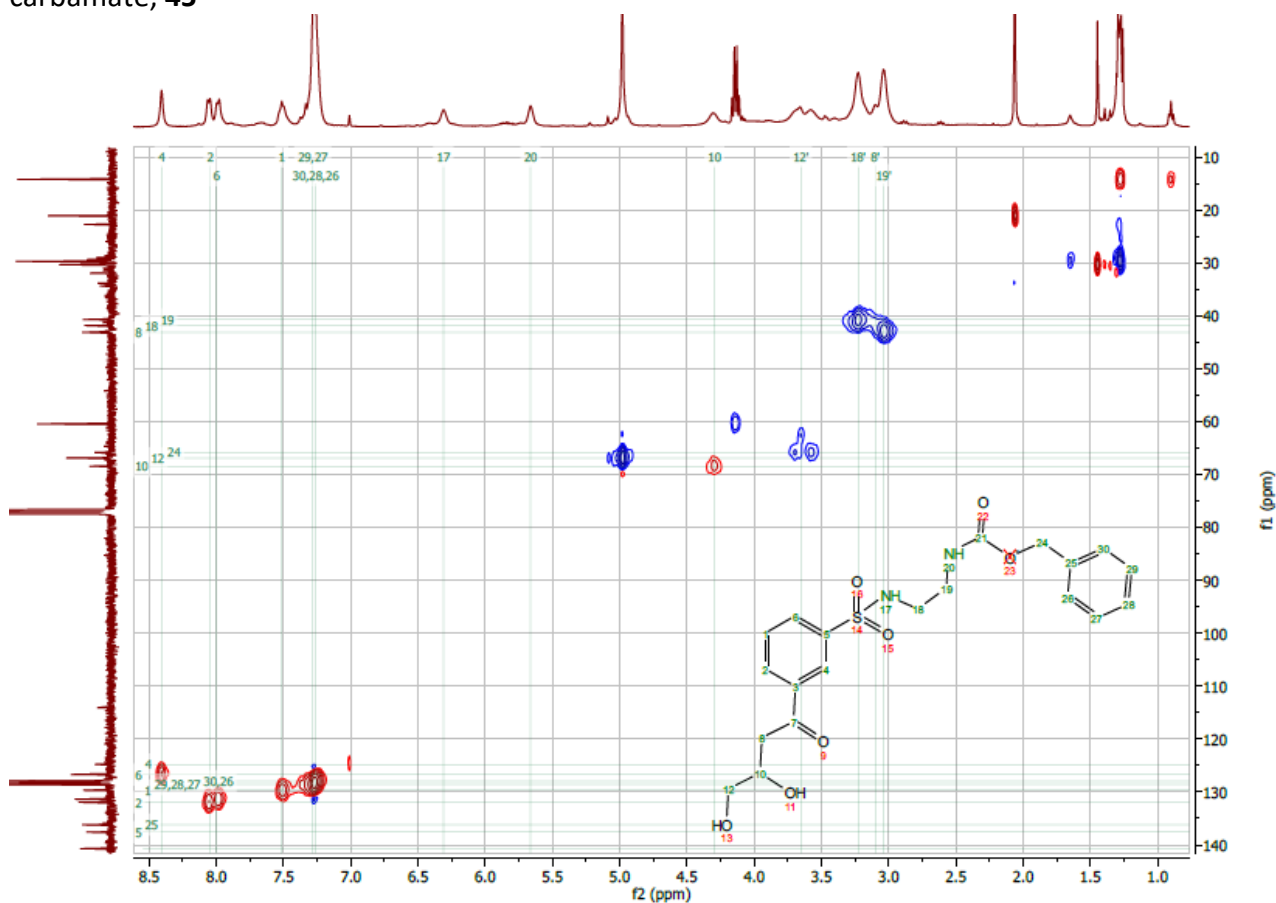
<sup>13</sup>C-NMR Benzyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl) carbamate, **45**



COSY-NMR Benzyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl) carbamate, **45**

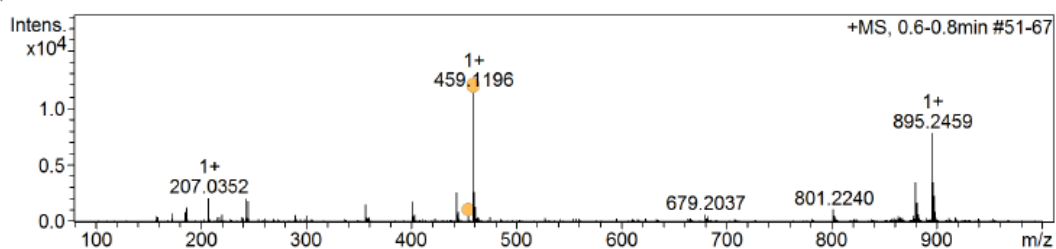


HSQC-NMR Benzyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **45**



ESI-MS Benzyl (2-((3-(3,4-dihydroxybutanoyl)phenyl)sulfonamido)ethyl)carbamate, **45**

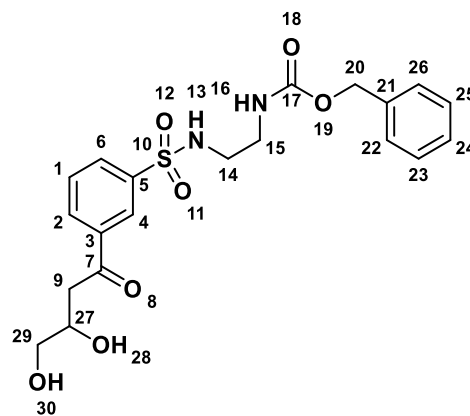
+MS, 0.6-0.8min #51-67



#	m/z	I %
1	185.1134	7.3
2	186.2207	11.0
3	207.0352	18.5
4	243.0742	18.1
5	245.0891	16.6
6	357.1261	13.1
7	401.1175	15.9
8	443.1248	23.0
9	445.1362	7.6
10	459.1196	100.0
11	460.1226	23.2
12	461.1249	11.3
13	801.2240	9.5
14	879.2521	31.3
15	880.2557	14.5
16	881.2613	15.1
17	895.2459	69.0
18	896.2499	31.2
19	897.2501	20.6
20	898.2538	7.6

← M+Na<sup>+</sup>

← 2M+Na<sup>+</sup>



[illegible]

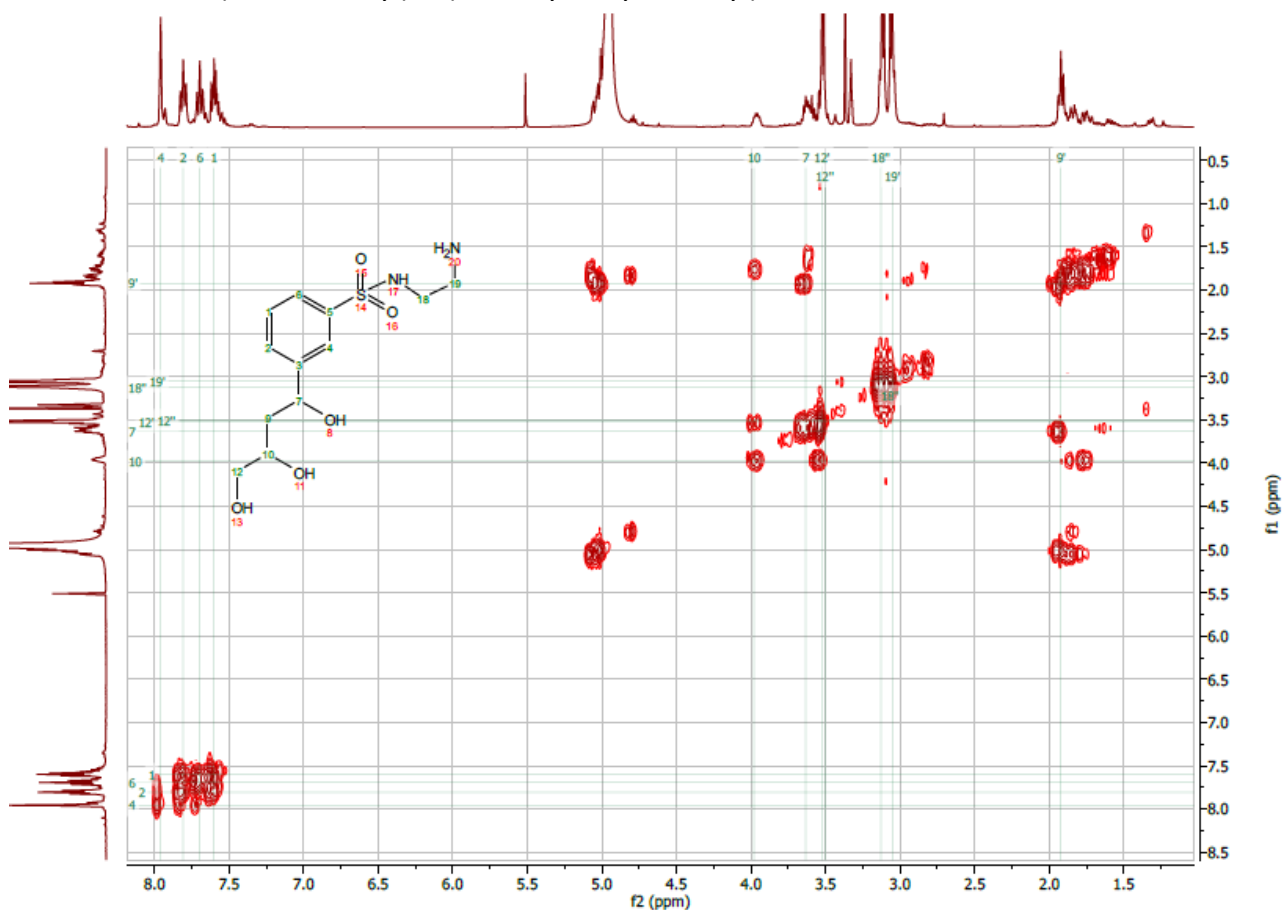
The figure displays the <sup>13</sup>C NMR spectrum of 2-amino-3-(hydroxymethyl)benzenesulfonamide. The chemical structure is shown with carbon atoms numbered 1 through 20. The spectrum features a series of peaks in the aromatic region (123-148 ppm), a cluster of peaks for the methoxy groups (53-55 ppm), a solvent peak for DMSO-d<sub>6</sub> at 40 ppm, and aliphatic peaks for the hydroxymethyl group (31-36 ppm). The x-axis represents the chemical shift in ppm, ranging from 25 to 145. The y-axis represents intensity, ranging from -2000 to 34000.

**Chemical Structure:** 2-amino-3-(hydroxymethyl)benzenesulfonamide. SMILES: Nc1cc(cc(c1CS(=O)(=O)N)CO)CO

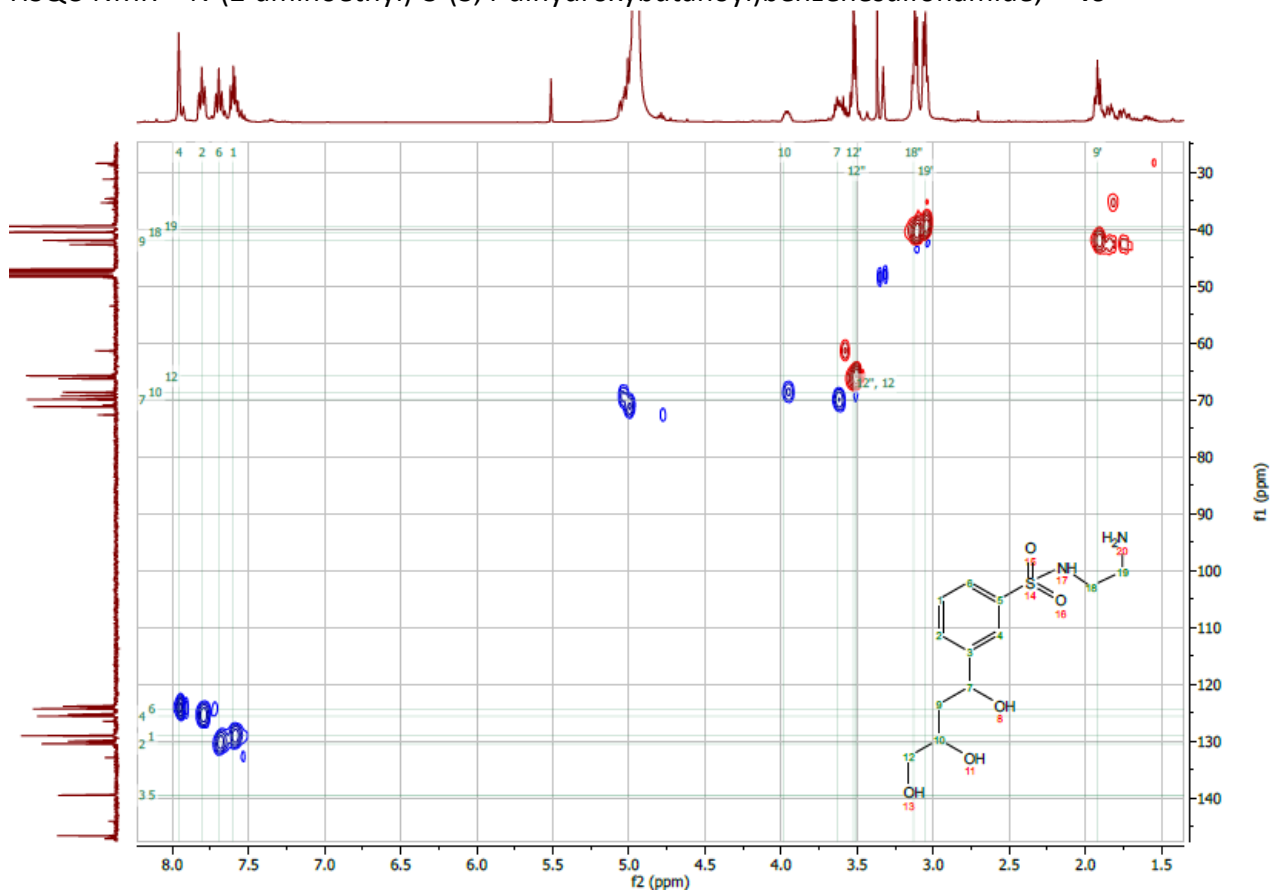
**Peak Assignments (ppm):**

- 147.76, 146.70, 144.12, 138.96, 138.54, 138.11, 137.46, 132.99, 130.50, 130.34, 130.05, 128.12, 128.09, 128.03, 126.96, 126.04, 125.13, 125.36, 124.27, 124.13, 123.90 (Aromatic and sulfonamide carbons)
- 72.65, 71.20, 70.92, 69.97, 69.35, 68.69, 66.29, 65.97, 65.75, 61.45, 61.37 (Solvent and water peaks)
- 53.49, 53.53 MeOD, 48.31 MeOD, 48.10 MeOD, 47.89 MeOD, 47.68 MeOD (Methoxy carbons)
- 40.00 (DMSO-d<sub>6</sub>)
- 35.36, 34.66, 31.21, 28.76, 28.41 (Aliphatic carbons)

COSY-NMR N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide, **46**

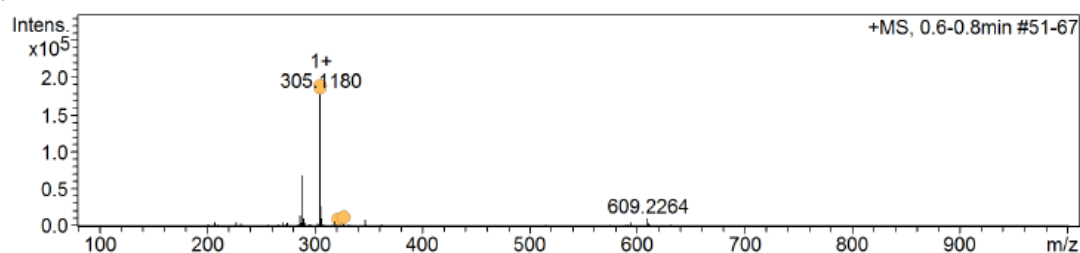


HSQC-NMR N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide, **46**



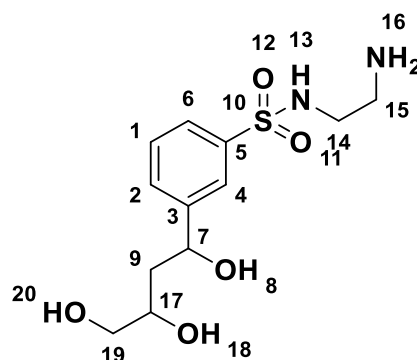
# ESI-MS N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide, **46**

+MS, 0.6-0.8min #51-67

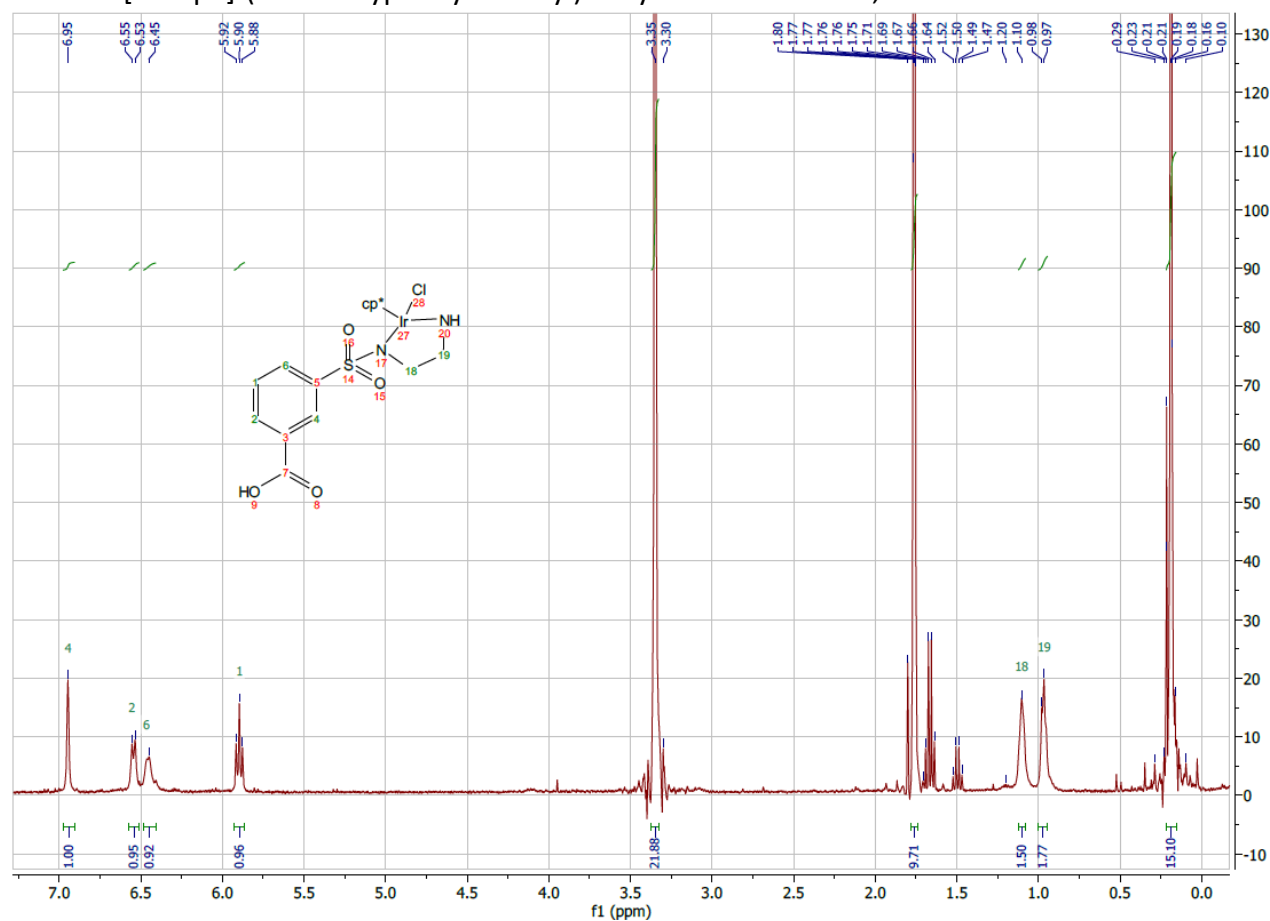


#	m/z	I %
1	207.0347	2.6
2	227.0907	2.9
3	231.0798	1.6
4	267.0808	1.6
5	271.1111	2.7
6	275.1062	2.1
7	287.1065	8.4
8	288.1107	2.2
9	289.1225	38.7
10	290.1241	5.7
11	291.1211	2.4
12	303.1301	1.6
13	305.1180	100.0
14	306.1202	15.7
15	307.1152	6.1
16	319.1328	3.5
17	347.1608	4.4
18	593.2321	2.7
19	609.2264	5.8
20	610.2277	1.8

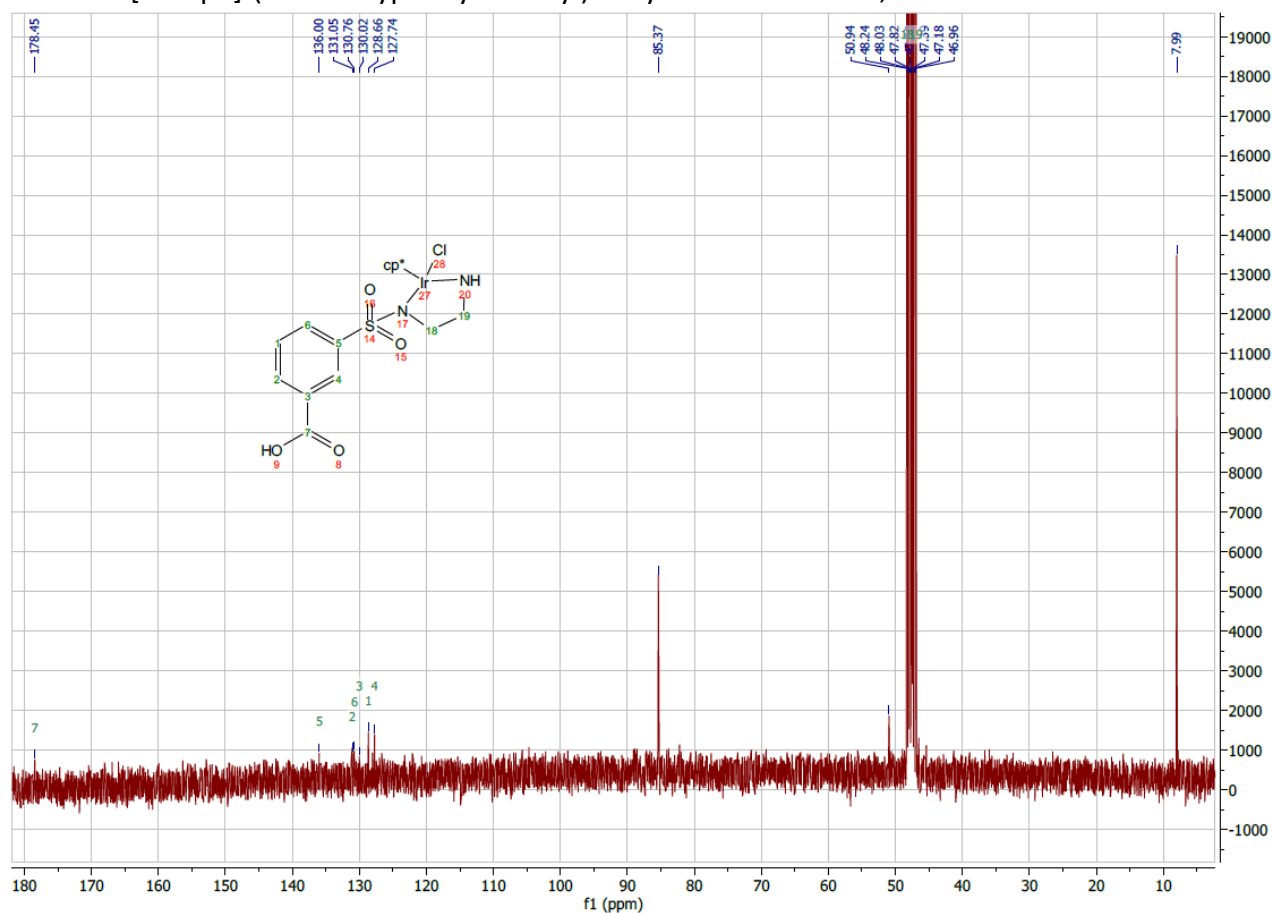
← M+H<sup>+</sup>



## <sup>1</sup>H-NMR [IrClCp\*]-(4-carboxyphenylsulfonyl)-ethylenediamine acid, **47**

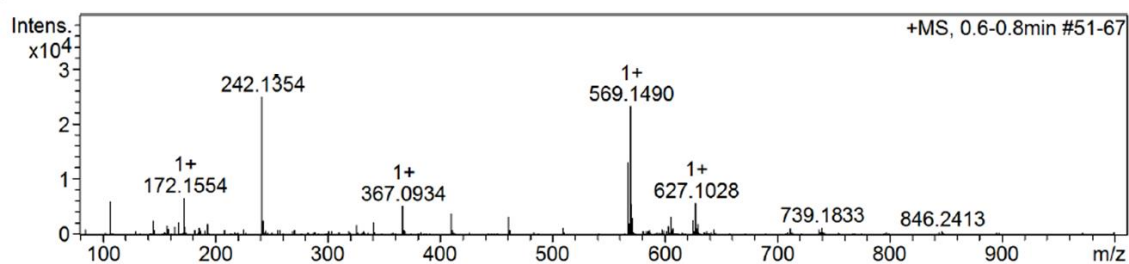


$^{13}\text{C}$ -NMR [IrClCp\*]-(4-carboxyphenylsulfonyl)-ethylenediamine acid, **47**



ESI-MS [IrClCp\*]-(4-carboxyphenylsulfonyl)-ethylenediamine acid, **47**

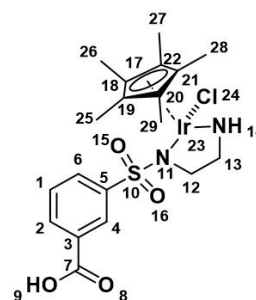
+MS, 0.6-0.8min #51-67



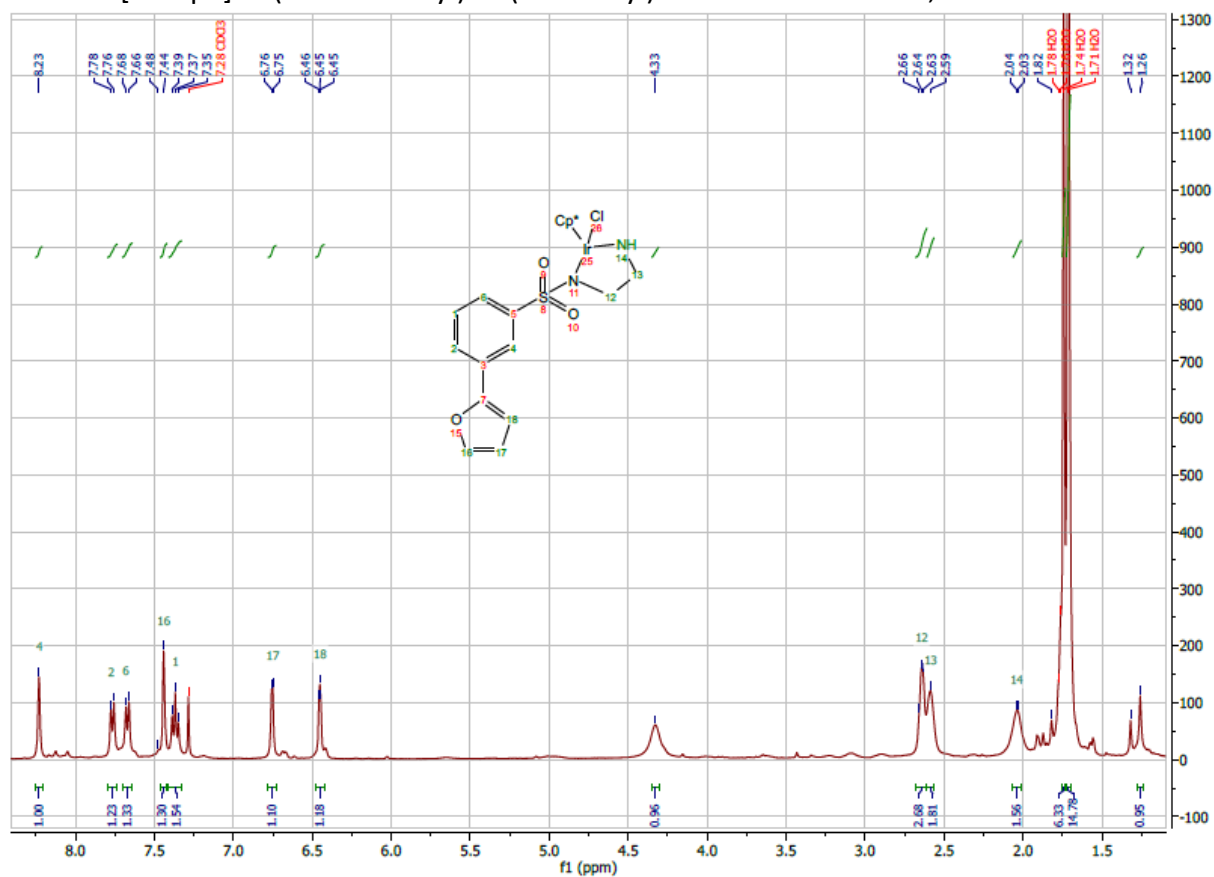
#	m/z	I %
1	107.0400	24.4
2	145.1328	10.2
3	167.1133	8.8
4	172.1554	27.1
5	193.1288	7.6
6	242.1354	100.0
7	243.1393	9.9
8	341.2174	8.7
9	367.0934	21.2
10	410.1331	15.7
11	461.2841	12.7
12	567.1395	52.3
13	568.1445	8.1
14	569.1490	93.3
15	570.1483	22.2
16	571.1487	11.8
17	605.1200	12.9
18	625.1041	10.3
19	627.1028	23.4
20	629.0994	7.8

[M- $^{37}\text{Cl}$ ] [M- $^{35}\text{Cl}$ ]

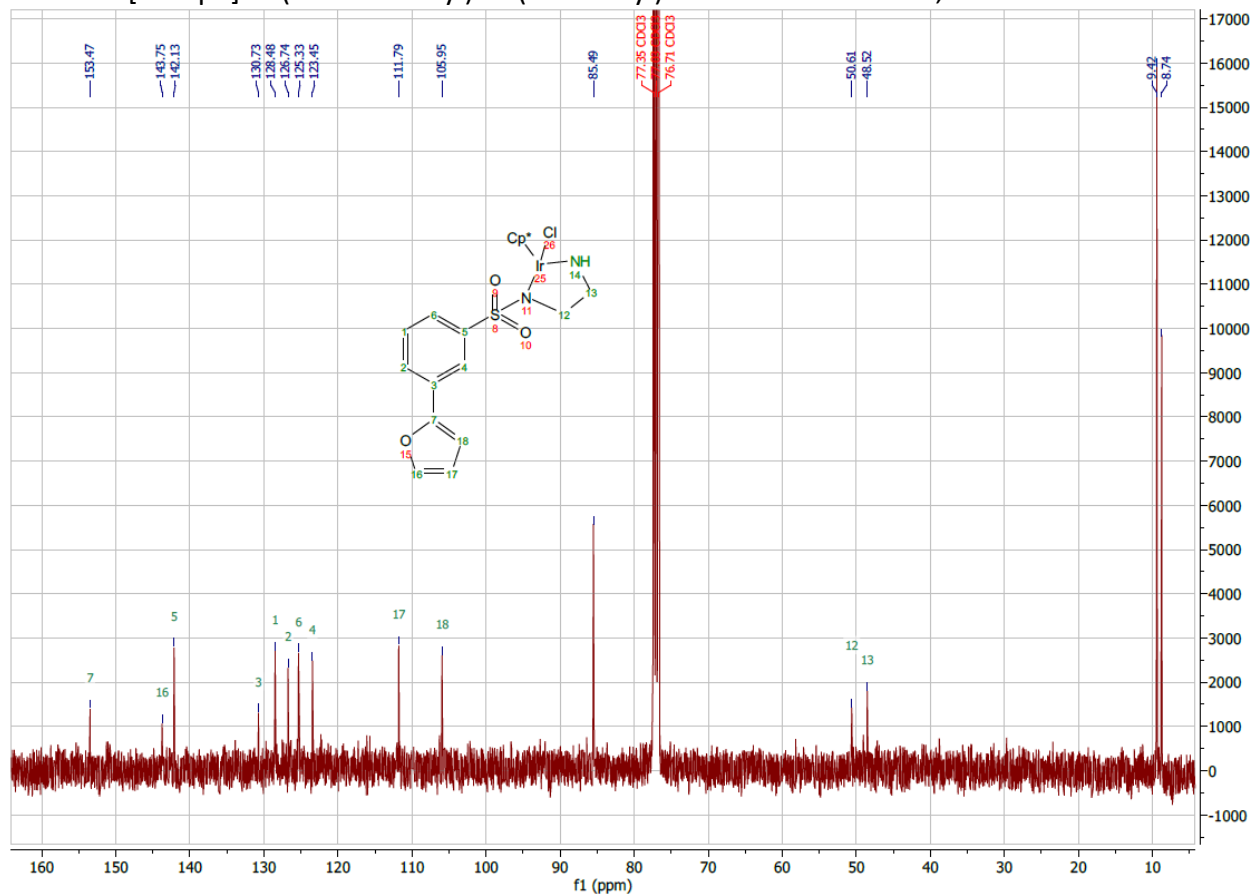
[M]



<sup>1</sup>H-NMR [IrClCp\*]-N-(2-aminoethyl)-3- (furan-2-yl)benzenesulfonamide, **49**



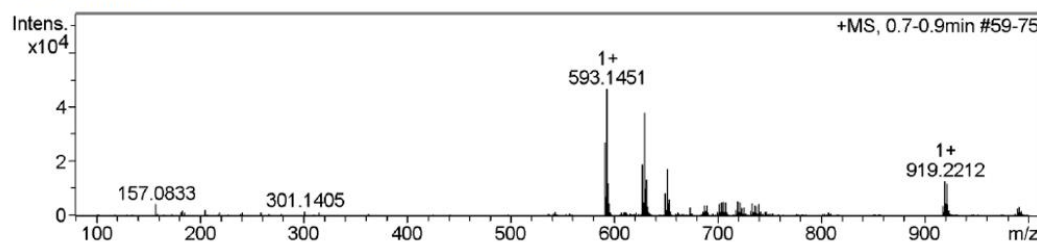
<sup>13</sup>C-NMR [IrClCp\*]-N-(2-aminoethyl)-3- (furan-2-yl)benzenesulfonamide, **49**





# ESI-MS [IrClCp\*]-N-(2-aminoethyl)-3- (furan-2-yl)benzenesulfonamide, **49**

+MS, 0.7-0.9min #59-75

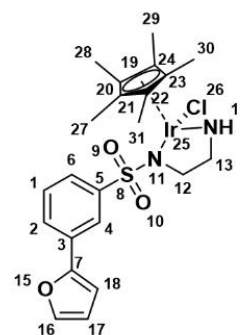


#	m/z	I %
1	591.1412	57.7
2	592.1445	14.9
3	593.1451	100.0
4	594.1482	25.9
5	627.1188	40.5
6	628.1234	10.4
7	629.1197	81.2
8	630.1222	21.3
9	631.1185	28.5
10	649.1002	17.8
11	651.1018	37.1
12	653.0990	12.7
13	703.1667	10.5
14	705.1520	10.8
15	707.1524	10.5
16	719.1786	11.5
17	721.1763	10.7
18	919.2212	27.3
19	920.2243	9.5
20	921.2225	25.5

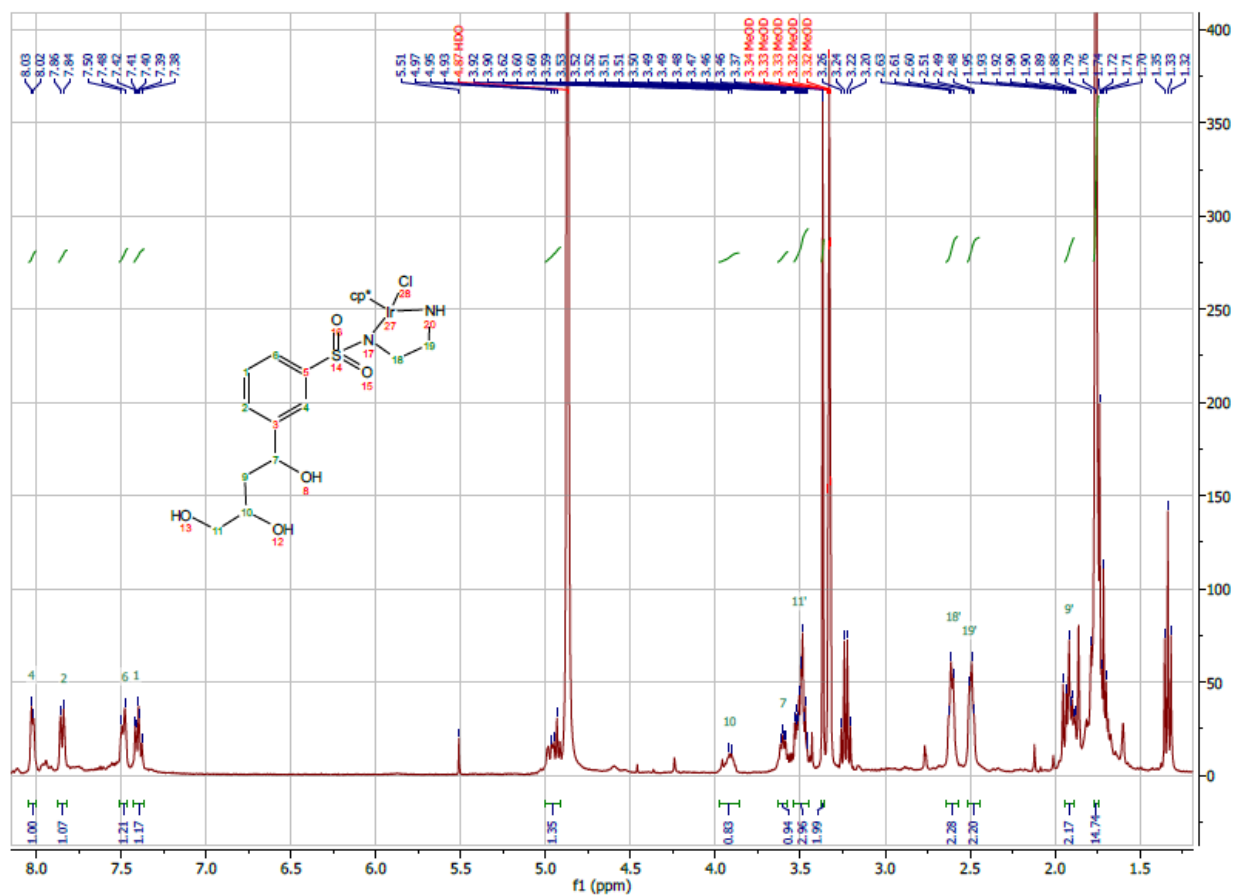
← [M+H<sup>+</sup>]-Cl

← [M+2H<sup>+</sup>]

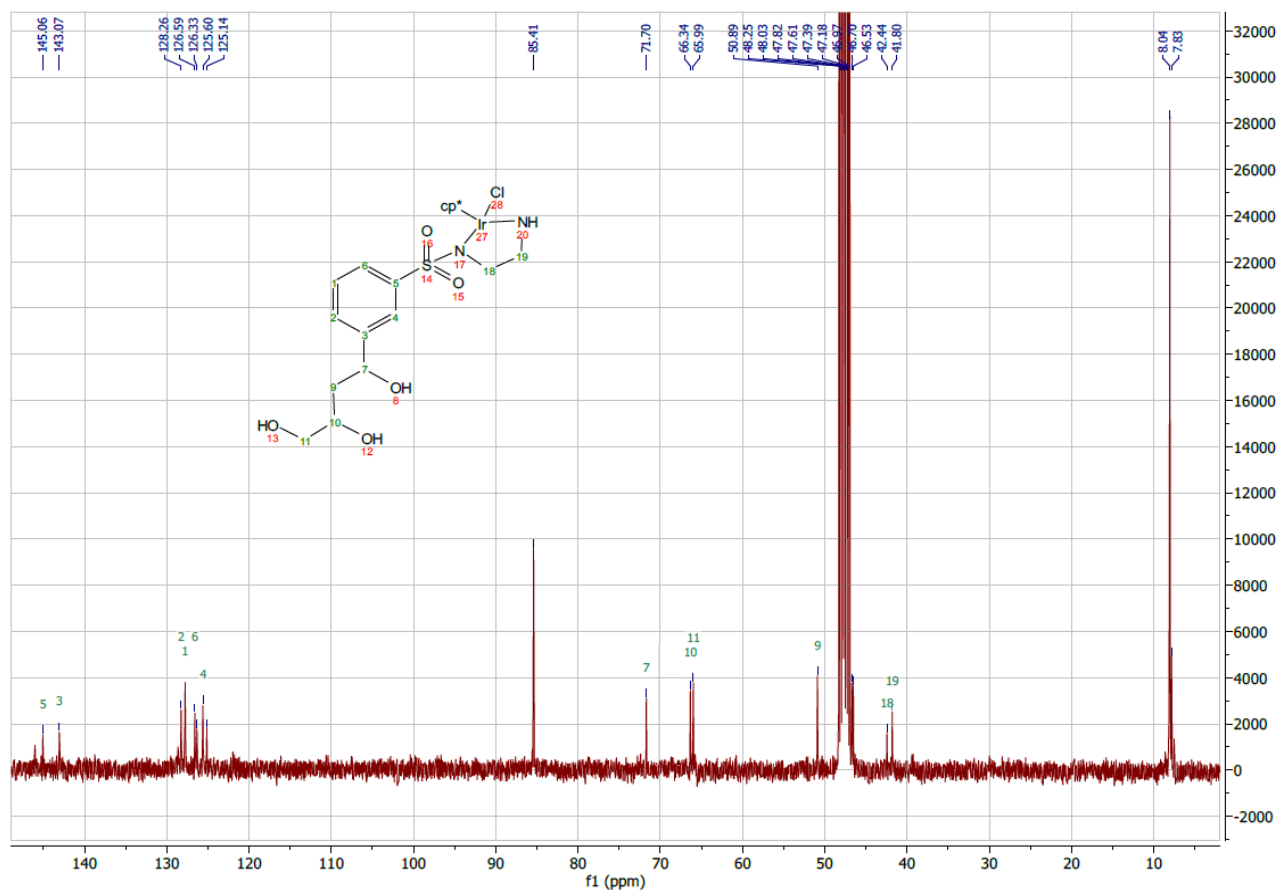
← [M+H<sup>+</sup>]+Na



## <sup>1</sup>H-NMR [IrClCp\*]- N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide, **50**

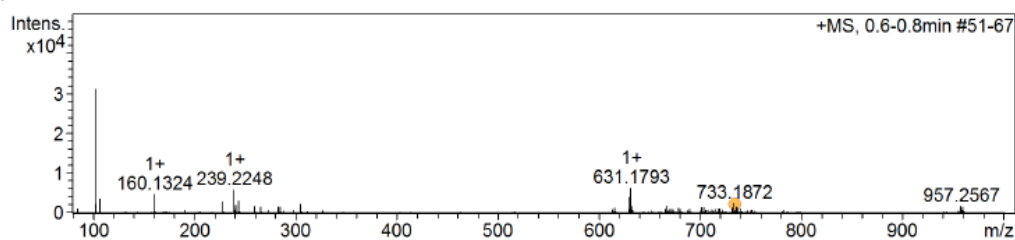


<sup>13</sup>C-NMR [IrClCp\*]-N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide,  
**50**



ESI-MS [IrClCp\*]- N-(2-aminoethyl)-3-(3,4-dihydroxybutanoyl)benzenesulfonamide,  
**50**

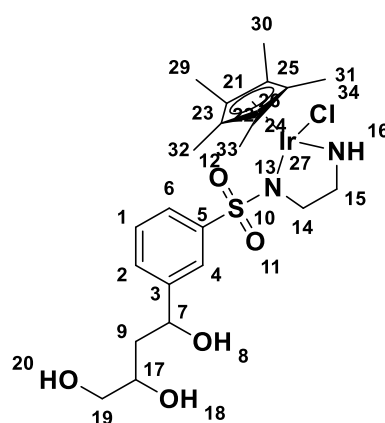
+MS, 0.6-0.8min #51-67



#	m/z	I %
1	102.1285	100.0
2	103.1309	7.4
3	107.0411	11.3
4	160.1324	15.3
5	228.0670	9.3
6	239.2248	19.0
7	241.2216	6.3
8	244.1540	10.0
9	260.0922	5.5
10	266.1360	5.2
11	285.1718	4.9
12	305.1159	7.2
13	629.1779	12.6
14	631.1793	20.3
15	632.1809	5.4
16	667.1558	6.0
17	733.1872	8.1
18	735.1849	5.1
19	957.2567	5.5
20	959.2558	4.9

← [M+H<sup>+</sup>]-Cl

← [M+H<sup>+</sup>]



## Appendix 2 - Horse Liver ADH, *Thermoanaerobacter brockii* ADH and *Thermoanaerobacter brockii* ADH mutant constructs

### His-HLADH and TbADH-WT

The HLADH gene inserted into the pRSETb expression vector was obtained as a gift from Dr Paradisi's group. The vector provides ampicillin (carbenicillin) resistance, and a C-terminus hexahistidine tag followed by an enterokinase cleavage site. The HLADH gene is cloned between the SacI and the EcoRI restriction sites, as displayed below with the start and stop codons highlighted in red, and the restriction sites highlighted in green.

```
ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGATC
TGTACGACGATGACGATAAGGATCCGAGCTCGATGAGCACAGCAGGAAAAGTAATAAAATGCAAAGCGGCT
GTGCTGTGGGAGGAAAAGAAACCATTTTCCATCGAGGAGGTGGAGTTGCACCCCGAAGGCCCATGAAGTC
CGTATAAAGATGGTGGCCACAGGAATTTGTCGCTCAGATGACCACGTGGTTAGTGGAACCTTGTCACACCTC
TTCTGTGATCGCAGGCCATGAGGCAGCGGGCATTGTGGAGAGCATTGGAGAAGGCGTCACTACAGTAAGAC
CAGGTGATAAAGTCATCCCACTTTACTCCCAAGTGTGGAAAATGCAGGGTTTGTAAAGCACCTGAAGGCAA
CTTCTGCTTGAAAAATGATCTGAGCATGCCTCGGGGAACCATGCAGGATGGTACCAGCAGGTTCACTGCAGA
GGGAAGCCCATCCCACTTCTTGGCACCAGCACCTTCTCCAGTACACCGTGGTGGACGAGATCTCAGTGG
CCAAGATCGATGCGGCCTCACCGCTGGAGAAAGTCTGTCTCATTGGCTGTGGATTTTCTACTGTTATGGGTCT
GCAGTCAAGGTTGCCAAGGTCACCCAGGGCTCCACCTGTGCCGTGTTGGCCTTGGAGGAGTGGGCCTGTCT
GTTATCATGGGCTGTAAAGCAGCCGGAGCGGCCAGGATCATTGGGGTGGACATCAACAAAGACAAGTTTGCA
AAGGCCAAAGAAGTGGGTGCCACTGAGTGTGTCAACCCTCAGGACTACAAGAAACCATCCAGGAGGTGCTG
ACAGAAATGAGCAATGGAGGTGTGGATTTTTCTTTGAAGTCATTGGTCGGCTCGACACTATGGTGAAGTGCCT
TGTCATGCTGTCAAGAAGCATATGGTGTGAGCGTCATTGTGGGAGTACCTCCTGATCCCAAATCTCTCTATG
AATCTATGTTGCTACTGAGTGGACGTACCTGGAAAGGAGCTATTTTGGCGGTTTTAAGAGTAAAGATTCTG
TCCCAAACTGTGGCCGATTTTATGGCTAAAAAGTTGCACTGGATCCTTTAATCACCCATGTTTTACCTTTG
AAAAAATAAATGAAGGATTTGACCTGCTTCGCTCTGGAGAGAGTATCCGTACCATCCTGACGTTTGAAGACCA
TACAAATGCTGCACCTGTAGCCGCTTCTGGCTCCTCTATCCTCTGGATCATCAGCCAAACGACATCAATAATT
CTGTTCTCAAAGATGCTATTAATAGTTACCGCTGGGAGCTTTCTAAAGAAACAAAAATTGATGTGAAGTCAC
TTTTCAAGCAAACGTTTAAATCCAAGTGAGAGCTAGAGGAACCATCAGCTGGGTAACTGAGCCCACTAACT
TTCTTCTTAATCATTCTCCTCACGTTGAATCCTGTACCTTTCCCGAATTC
```

The His-HLADH primary sequence is shown below. The C-terminus hexahistidine tag and enterokinase cleavage site are listed as bold letters. Predicted properties: MW = 43850; pI = 7.5.

```
MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPSSMSTAGKVIKCKAAVLWEEKKPFSEIEVEVAPPKAHEV
RIKMVATGICRSDHVVSGTLVPLPVIAGHEAAGIVESIGEGVTVRPGDKVIPLFTPQCGKCRVCKHPEGNFCLK
NDLSMPRGTMQDGTSRFTCRGKPIHHFLGTSTFSQYTVVDEISVAKIDAASPLEKVCCLIGCGFSTGYGSAVKVAVT
QGSTCAVFLGGVGLSVIMGCKAAGAARIIGVDINKDKFAKAKEVGATECVNPQDYKKPIQEVLTMSNGGVDFS
FEVIGRLDTMVTALSCCQEAYGVSIVGVPPDSQNLNLMNPMLLSGRTWKGAIFGGFKSKDSVPKLVADFMAKKF
ALDPLITHVLPFEKINEGFDLLRSGESIRTILTF
```

The TbADH-WT gene fused to an N-terminus Strep-tag (II) followed by a thrombin cleavage site was synthesized by Biomatik (Ontario, Canada) and inserted into the pET-21a(+) expression vector. The vector provides ampicillin (carbenicillin) resistance. The TbADH-WT gene is cloned between the NdeI and the XhoI restriction sites, as presented below with the start and stop codons highlighted in red, and the restriction sites highlighted in green.

```
AGAAGCTAACTCCCTCTAGAATAATTTGTTTAACTTTAAGAAGGAGATATACATATGTGGAGCCACCCGCA
GTTCGAAAAATCTTCTGGTCTGGTCCGCGTGGATCCATGAAGGGCTTCGCGATGCTGAGCATTGGTAAGGTT
```

GGCTGGATCGAGAAAGAAAAACCGGCTCCGGGTCCGTTTCGACGCTATCGTGCGTCCGCTGGCTGTTGCGCCG  
TGCACCTCTGACATCCATACCGTTTTTCGAGGGTGCGATCGGTGAGCGCCACAACATGATCCTGGGTACGAAAG  
CGGTGGGCGAGGTTGTTGAAGTTGGCTCTGAGGTGAAAGACTTTAAGCCGGGTGATCGTGTTGTTGTTCCGG  
CAATACCCCGGACTGGCGTACTTCCGAAGTTCAGCGCGGCTACCACCAGCACTCTGGCGGTATGCTGGCGGG  
TTGGAAATTCTCCAACGTTAAGGACGGCGTGTTCCGCGAGTTCTTCCATGTGAACGACGCTGACATGAACCTG  
GCGCACCTGCCGAAGGAAATCCCGCTGGAAGCGGCGGTTATGATTCCGGATATGATGACTACTGGTTTCCATG  
GTGCAGAGCTGGCAGACATTGAGCTGGGTGCAACCGTTGCGGTTCTGGGTATCGGTCCGGTTGGCCTGATGG  
CGGTGGCTGGTGCGAAACTGCGTGGTGCGGGTTCGTATCATCGCGGTTGGTTCTCGTCCGGTGTGCGTTGATG  
CTGCTAAATACTACGGTGCGACCGACATTGTTAACTACAAAGACGGTCCGATCGAATCTCAGATTATGAACCT  
GACCGAGGGCAAAGGCGTGGACGCGAGCGATTATCGCAGGTGGTAACGCGGATATCATGGCGACCGCTGTTA  
AAATCGTTAAACCGGGTGGTACTATTGCGAACGTGAACTATTTCCGTGAAGGCGAAGTGCTGCCGGTGCCGC  
GTCTGGAATGGGGTTGTGGTATGGCACATAAGACCATTAAAGGTGGTCTGTGTCCGGGCGGTCTGTGCGCA  
TGGAACGTCTGATTGACCTGGTTTTCTACAAACGTGTTGACCCGTCTAAACTGGTTACCCACGTGTTCCGTGGT  
TTCGACAACATCGAAAAGGCTTTTATGCTGATGAAAGATAAACCGAAAGATCTGATCAAACCGGTGGTTATCC  
TGCGCTAACTCGAGCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAAAGAAGCGTACGTC

The TbADH-WT primary sequence is shown below. The *N*-terminus Strep-tag (II) and thrombin cleavage site are listed as bold letters. Predicted properties: MW = 39659.18; pI = 6.61.

MWSHPQFEKSSGLVPRGSMKGFAMLSIGKVGWIEKEKPAPGPFDAIVRPLAVAPCTSDIHTVFEGAIGERHNMIL  
GHEAVGEVVEVGSEVKDFKPGDRVVPAITPDWRTSEVQRGYHQHSGGMLAGWKFSNVKDGVFGEFFHVND  
DMNLAHLPEIPLEAAVMIPDMMTTGFHGAELADIELGATVAVLGIGPVGLMAVAGAKLRGAGRII  
DAAKYYGATDIVNYKDGPIESQIMNLTEGKGVDAIIAGGNADIMATAVKIVKPGGTIANVNYFGE  
GVEVLPVPRLEWGCGMAHKTIKGLCPGRLRMRERLIDLVFYKRVDPKLVTHVFRGFDNIEKAFMLMKDKPKDLIKPVVILA

## TbADH-5M

The mutated TbADH-5M gene was obtained from Dr. Mattias Basle. The DNA sequence of the gene is displayed below. The different features are highlighted as follow: restriction sites in green, start and stop codons in blue, the *N*-terminus strep-tag (II) in brown, the mutations applied to gene in red.

AGAAGCTAACTCCCTCTAGAATAATTTGTTTAACTTTAAGAAGGAGATATACATATGTGGAGCCACCCGCA  
GTTTCGAAAACTCTTCTGGTCTGGTTCGCGTGGATCCATGCTGAGCATTGGTAAGGTT  
GGCTGGATCGAGAAAGAAAAACCGGCTCCGGGTCCGTTTCGACGCTATCGTGCGTCCGCTGGCTGTTGCGCCG  
TGCACCTCTGACATCCATACCGTTTTTCGAGGGTGCGATCGGTGAGCGCCACAACATGATCCTGGGTGCCGAAG  
CGGTGGGCGAGGTTGTTGAAGTTGGCTCTGAGGTGAAAGACTTTAAGCCGGGTGATCGTGTTGTTGTTCCGG  
CAATACCCCGGACTGGCGTACTTCCGAAGTTCAGCGCGGCTACCACCAGCACTCTGGCGGTATGCTGGCGGG  
TTGGAAATTCTCCAACGTTAAGGACGGCGTGTTCCGCGAGTTCTTCCATGTGAACGACGCTGACATGAACCTG  
GCGCACCTGCCGAAGGAAATCCCGCTGGAAGCGGCGGTTATGATTCCGGCTATGATGACTACTGGTTTCCATG  
GTGCAGAGCTGGCAGACATTGAGCTGGGTGCAACCGTTGCGGTTCTGGGTATCGGTCCGGTTGGCCTGATGG  
CGGTGGCTGGTGCGAAACTGCGTGGTGCGGGTTCGTATCATCGCGGTTGGTTCTCGTCCGGTGAGCGTTGATG  
CTGCTAAATACTACGGTGCGACCGACATTGTTAACTACAAAGACGGTCCGATCGAATCTCAGATTATGAACCT  
GACCGAGGGCAAAGGCGTGGACGCGAGCGATTATCGCAGGTGGTAACGCGGATATCATGGCGACCGCTGTTA  
AAATCGTTAAACCGGGTGGTACTATTGCGAACGTGAACTATTTCCGTGAAGGCGAAGTGCTGCCGGTGCCGC  
GTCTGGAATGGGGTGCTGGTATGGCACATAAGACCATTAAAGGTGGTCTGGCTCCGGGCGGTCTGTGCGCA  
TGGAACGTCTGATTGACCTGGTTTTCTACAAACGTGTTGACCCGTCTAAACTGGTTACCCACGTGTTCCGTGGT  
TTCGACAACATCGAAAAGGCTTTTATGCTGATGAAAGATAAACCGAAAGATCTGATCAAACCGGTGGTTATCC  
TGCGCTAACTCGAGCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAAAGAAGCGTACGTC  
C

### Appendix 3 – Kinetics Michaelis-Menten of His-HLADH oxidative reaction.

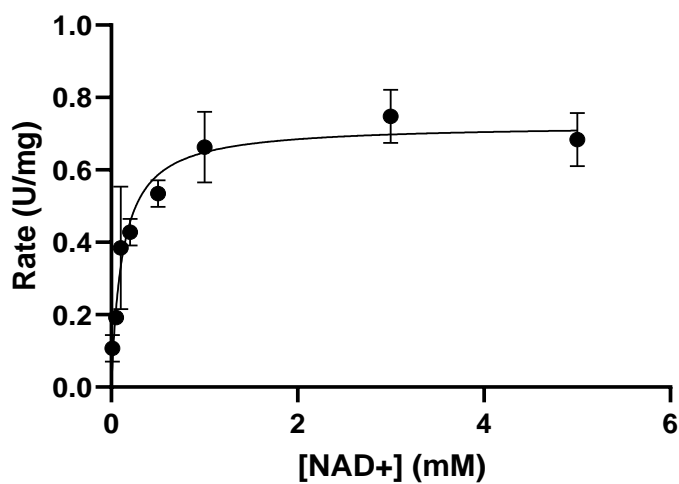


Figure 89: Michaelis-Menten curve of HLADH in the oxidation of ethanol.

### Appendix 4 – Inhibition kinetics graphs for His-HLADH and TbADH-WT.

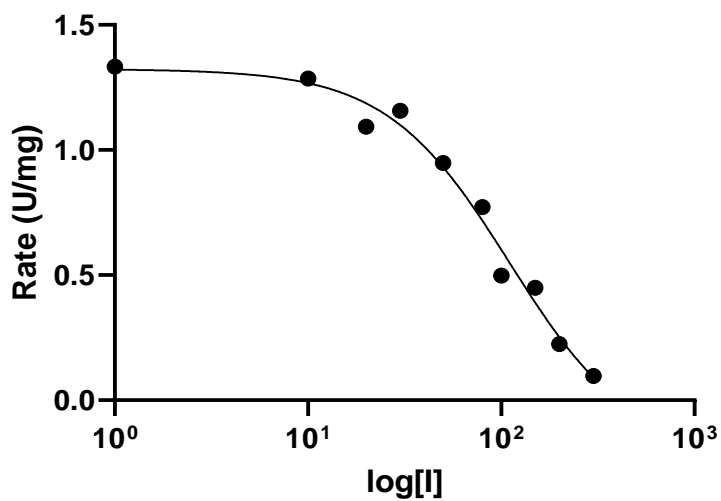


Figure 90: Inhibition test on His-HLADH with anchor benzoic acid 36.

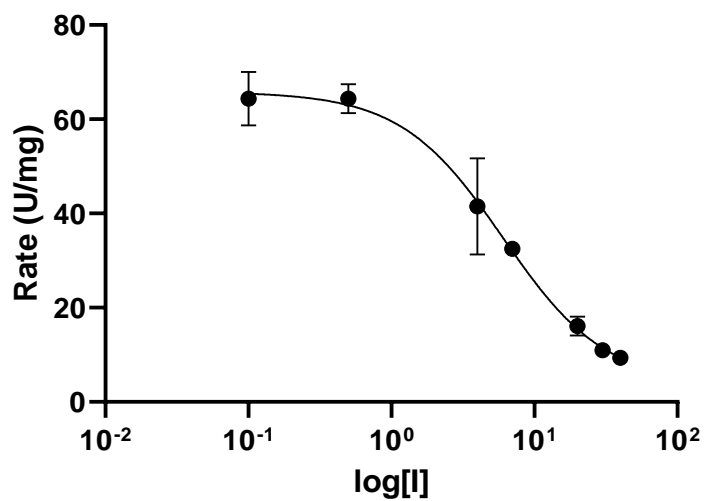


Figure 91: Inhibition test on TbADH-WT with anchor benzoic acid 36.

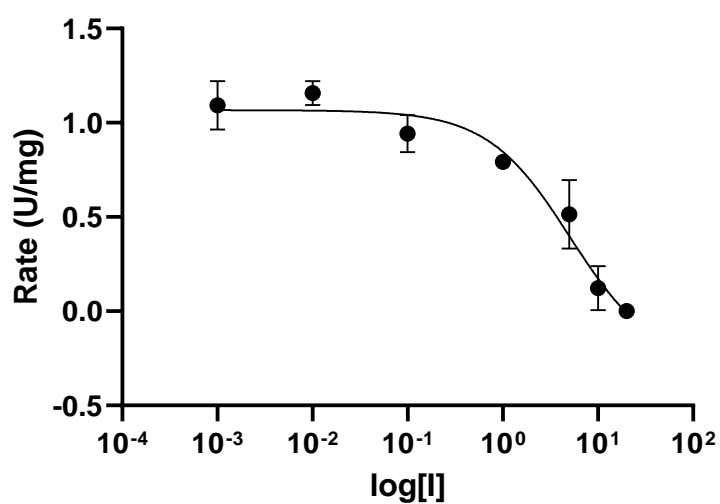


Figure 92: Inhibition test on His-HLADH with anchor furan 44.

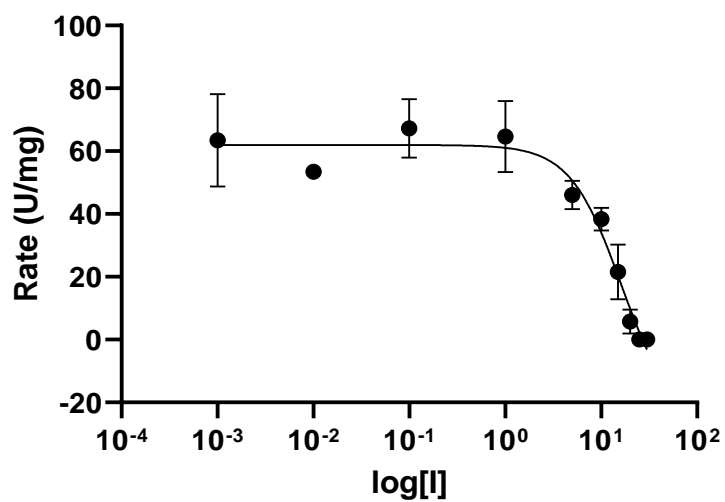


Figure 93: Inhibition test on TbADH-WT with anchor furan 44.

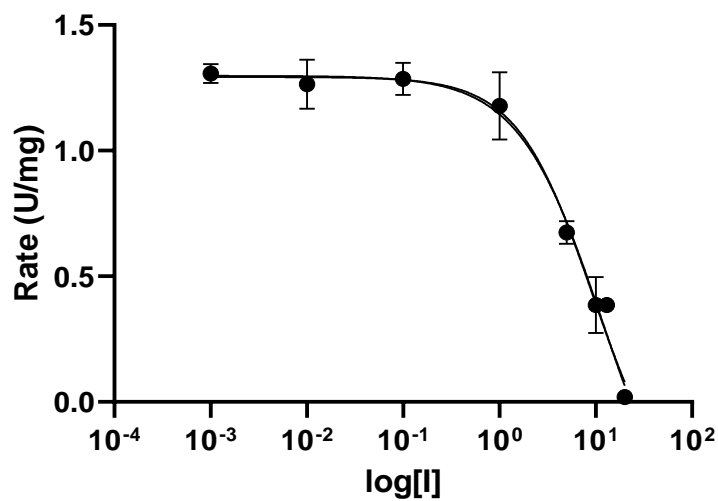


Figure 94: Inhibition test on His-HLADH with anchor Triol 46.

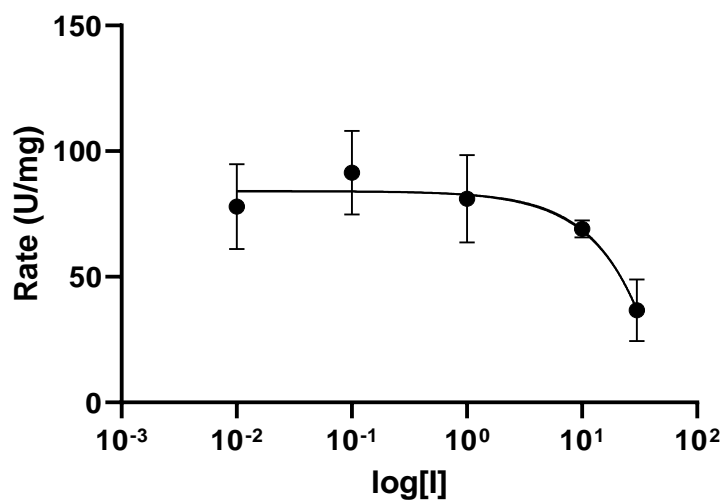


Figure 95: Inhibition test on TbADH-WT with anchor Triol 46.

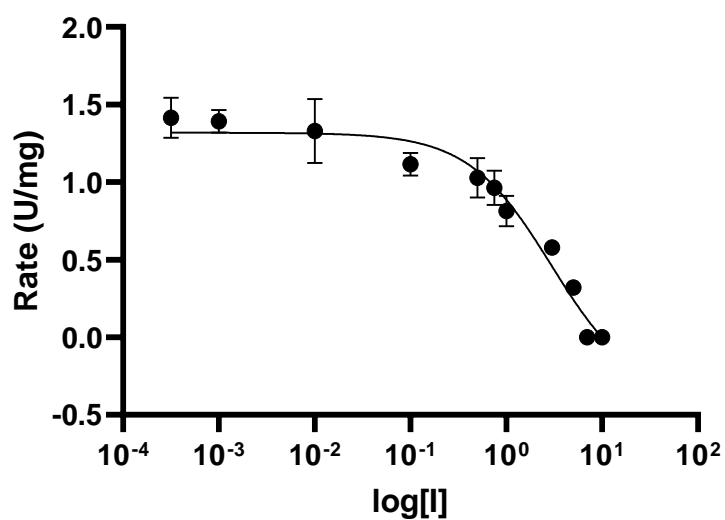


Figure 96: Inhibition test on His-HLADH with anchor [IrClCp\*]-benzoic acid 47.

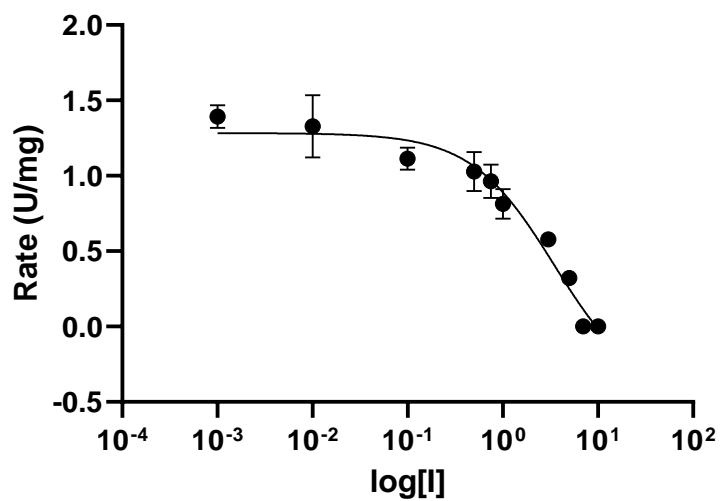


Figure 97: Inhibition test on TbADH-WT with anchor [IrClCp\*]-benzoic acid **47**.

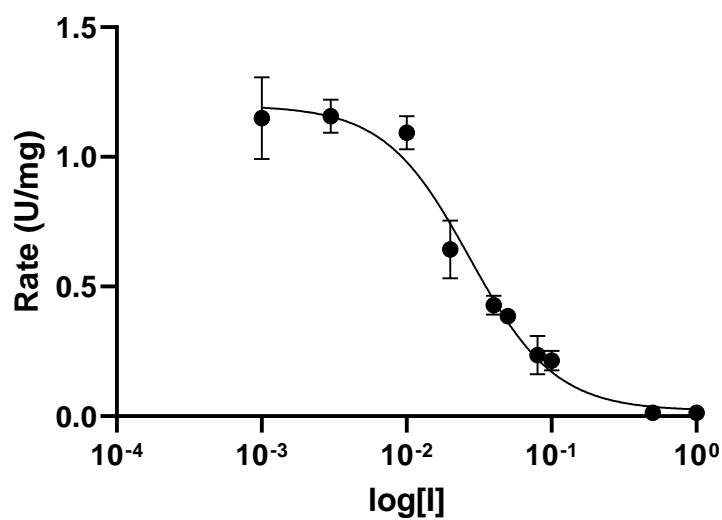


Figure 98: His-HLADH inhibition test with anchor [IrClCp\*]-furan **49**.

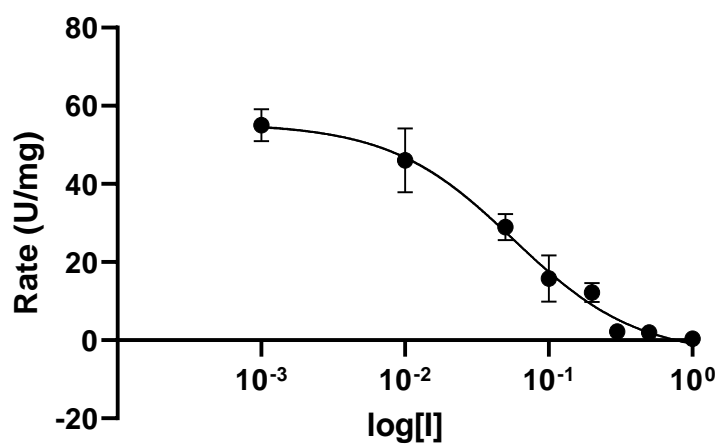


Figure 99: Inhibition test on TbADH-WT with anchor [IrClCp\*]-furan **49**.



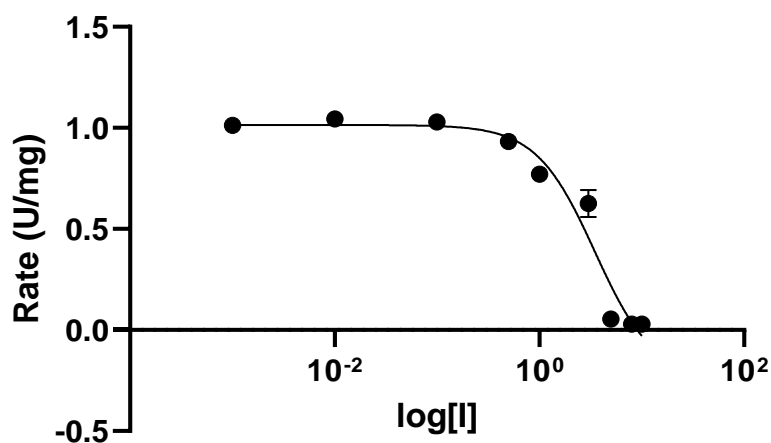


Figure 100: Inhibition test on His-HLADH with anchor [IrClCp\*]-triol 50.

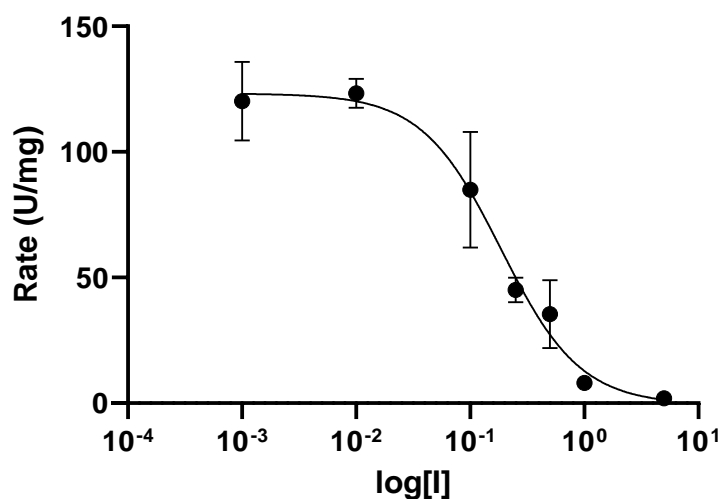


Figure 101: Inhibition test on TbADH-WT with anchor [IrClCp\*]-triol 50.

## Appendix 5 – Catalysis assays.

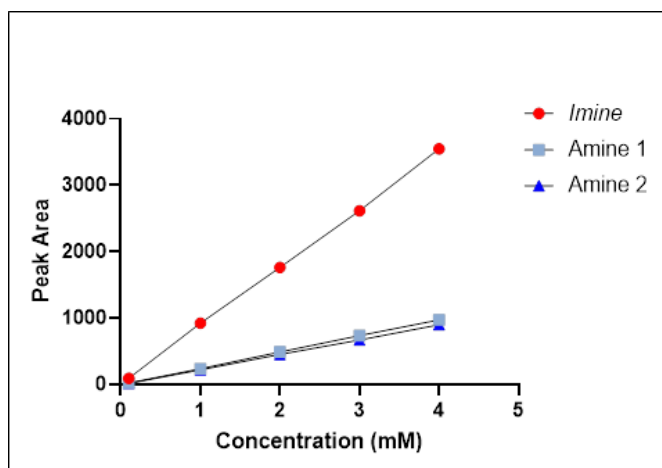
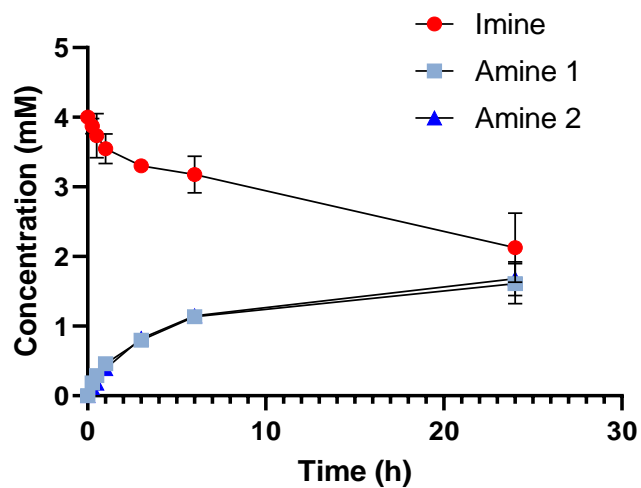
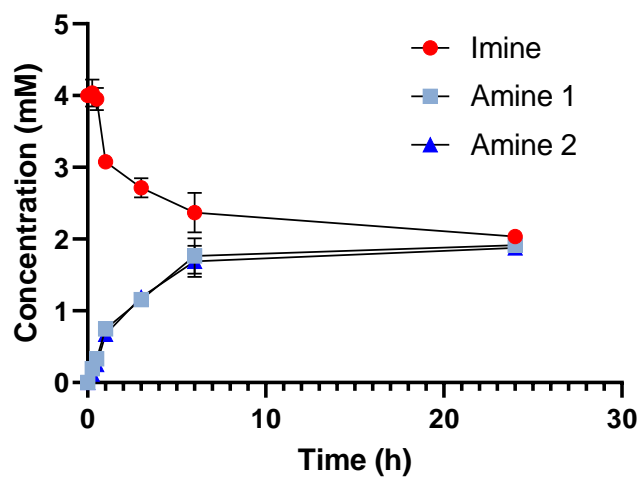


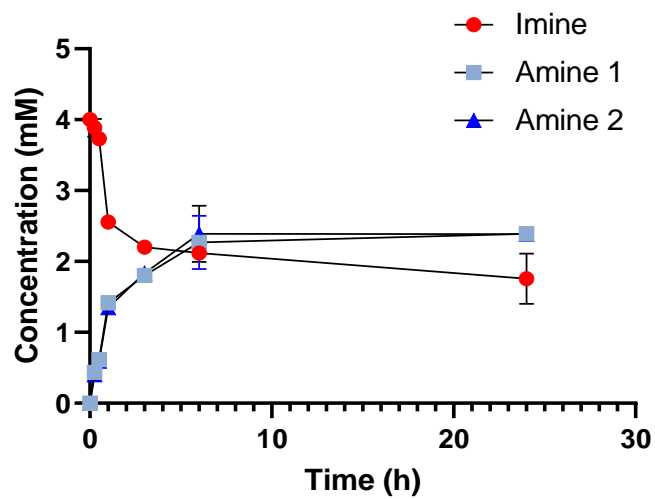
Figure 102: Calibration curve of the imine 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline substrate and the two chiral amine 6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline products (Amine 1 and 2).



**Figure 103:** The kinetic profile for the reduction of 4mM 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline at 38 °C catalysed by 50  $\mu$ M iridium complex [IrClCp\*]-benzoic acid **47**.



**Figure 104:** The kinetic profile for the reduction of 4mM 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline at 38 °C catalysed by 50  $\mu$ M iridium complex [IrClCp\*]-furan **49**.



**Figure 105:** The kinetic profile for the reduction of 4mM 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline at 38 °C catalysed by 50  $\mu$ M iridium complex [IrClCp\*]-triol **50**.